Indiana Law Review

Volume 36 No. 1 2003

Raise the Bar

BNA's *United States Law Week*

For 65 years, *U.S. Law Week* has consistently raised the standard of excellence in legal reporting. Today's *Law Week* is a powerful research and notification tool that does more for you than ever:

- **Case Alert:** A summary and analysis of the most significant court opinions nationwide, with electronic access to full text.
- **Legal News:** Reports on non-judicial developments, including legal analysis and perspective, important legislation, and administrative rulemakings.
- **Supreme Court Today:** Daily coverage of the status of every case before the U.S. Supreme Court and all Court proceedings.
- **Supreme Court Opinions:** Full text of decisions, with headnotes and counsel listings, daily online, weekly in print.

There's no more complete, authoritative way to keep up with all areas of the law than *U.S. Law Week.*

FREE TRIALS!
In print for 5 weeks FREE! **CALL 800-372-1033**
On the Web for 15 days FREE!
GO TO www.bna.com/trials/lwex511.htm

Visit BNA's Professional Information Center
http://litigationcenter.bna.com

BNA®

Essential information.
Expert analysis.

Indiana Law Review

Volume 36 2002-2003

Indiana Law Review

Volume 36 **2003** **Number 1**

TABLE OF CONTENTS

TRIBUTES

ARTICLES

NOTES

INDIANA LAW REVIEW

(ISSN 0090-4198)

Published four times a year by Indiana University. Editorial and Business Offices are located at:

Indiana Law Review
Lawrence W. Inlow Hall
530 W. New York Street
Indianapolis, IN 46202-3225
(317) 274-4440

Subscriptions. Current subscription rates for an academic year are $30.00 (domestic mailing) and $35.00 (foreign mailing) for four issues. Unless the Business Office receives notice to the contrary, all subscriptions will be renewed automatically. ***Address changes must be received at least one month prior to publication to ensure prompt delivery and must include old and new address and the proper zip code.***

Single and Back Issues. Current single regular issues may be obtained from the *Indiana Law Review* Business Office for $10.00. Current Survey issues may be purchased for $20.00. The Symposium Issue may be purchased for $15.00. Please enclose payment with order. Claims for nonreceipt of current year's issues must be made within six months of the mailing date. Back issues may be ordered from: William S. Hein & Co., Inc., 1285 Main St., Buffalo, NY 14209-1987, (800) 828-7571, (716) 883-8100 fax, mail@wshein.com (e-mail for orders).

Manuscripts. The *Indiana Law Review* invites submission of unsolicited manuscripts. All citations should conform to *The Bluebook A Uniform System of Citation* (17th ed. 2000). Manuscripts should be double-spaced with standard one-inch margins. The *Indiana Law Review*, as a matter of policy, promotes the use of gender-neutral language, and alternate use of the female or male pronoun is intended to refer to both female and male.

The INDIANA LAW REVIEW (ISSN 0090-4198) is the property of Indiana University and is published quarterly by the Indiana University School of Law—Indianapolis, which assumes complete editorial responsibility thereof. Current subscription rates for an academic year are: one year domestic $30.00; foreign $35.00. Send all correspondence to Editorial Specialist, *Indiana Law Review*, Indiana University School of Law—Indianapolis, Lawrence W. Inlow Hall, 530 W. New York Street, Indianapolis, Indiana 46202-3225. Publication office: 530 W. New York Street, Indianapolis, Indiana 46202-3225. Periodicals postage paid at Indianapolis, Indiana 46201.

POSTMASTER: Send address changes to INDIANA LAW REVIEW, Lawrence W. Inlow Hall, 530 W. New York Street, Indianapolis, Indiana 46202-3225.

 The entire text of this Law Review is printed on recycled paper.

Please enter my subscription
to the
INDIANA LAW REVIEW

NAME _____

ADDRESS: _____

_____

_____

Enclosed is $ _____ for _____ subscription(s)

Mail to: **ATTN: Law Review Specialist**
at
INDIANA LAW REVIEW
INDIANA UNIVERSITY SCHOOL OF LAW—**INDIANAPOLIS**
Lawrence W. Inlow Hall
530 West New York Street
Indianapolis, Indiana 46202-3225

For an *academic* year, the subscription rate for four issues is:
Domestic, $30; Foreign, $35; Student, $20
Single Issue, $10; Survey Issue, $20
Symposium Issue, $15

2002-2003 ADMINISTRATIVE OFFICERS AND FACULTY

Administrative Officers

GERALD L. BEPKO, *Interim President, Indiana University and Professor of Law.* B.S., Northern Illinois University; J.D., ITT/Chicago-Kent College of Law; LL.M., Yale University.

WILLIAM M. PLATER, *Interim Chancellor, Indiana University-Purdue University—Indianapolis.* B.A., M.A., Ph.D., University of Illinois.

ANTHONY A. TARR, *Dean and Professor of Law.* B.A., LL.B., University of Natal; LL.M., Cambridge University; Ph.D., University of Canterbury; Ph.D., Cambridge University.

THOMAS B. ALLINGTON, *Associate Dean for Technology and Professor of Law.* B.S., J.D., University of Nebraska; LL.M., New York University.

JEFFREY W. GROVE, *Associate Dean for Graduate Studies, Professor of Law, and Director, China Summer Program.* A.B., Juniata College; J.D., George Washington University.

SUSANAH M. MEAD, *Associate Dean for Academic Affairs and Professor of Law.* B.A., Smith College; J.D., Indiana University—Indianapolis.

CYNTHIA BAKER, *Director, Program on Law and State Government.* B.A., J.D., Valparaiso University.

ELIZABETH L. DECOUX, *Assistant Dean for Student Affairs.* J.D., Mississippi College School of Law.

ANGELA M. ESPADA, *Assistant Dean for Admissions.* J.D., Indiana University—Bloomington.

JONNA KANE MACDOUGALL, *Assistant Dean for External Affairs.* J.D., Indiana University—Indianapolis.

CAROL B. NEARY, *Director of Development.* B.A., Indiana University—Indianapolis.

SHANNON L. WILLIAMS, *Director of Professional Development.* B.S., Indiana University—Indianapolis.

Faculty

CYNTHIA M. ADAMS, *Clinical Associate Professor of Law.* B.A., Kentucky Wesleyan College; J.D., Indiana University—Indianapolis.

THOMAS B. ALLINGTON, *Associate Dean for Technology and Professor of Law.* B.S., J.D., University of Nebraska; LL.M., New York University.

JAMES F. BAILEY, III, *Professor of Law and Director of Law Library.* A.B., J.D., M.A.L.S., University of Michigan.

GERALD L. BEPKO, *Interim President, Indiana University and Professor of Law.* B.S., Northern Illinois University; J.D., ITT/Chicago-Kent College of Law; LL.M., Yale University.

FRANK BOWMAN, *Associate Professor of Law.* B.A., Colorado College; J.D., Harvard Law School.

WILLIAM C. BRADFORD, *Assistant Professor of Law.* B.A., M.A., University of Miami; Ph.D., Northwestern University; J.D., University of Miami School of Law; LL.M., Harvard Law School.

ROBERT BROOKINS, *Professor of Law.* B.S., University of South Florida; J.D., Ph.D., Cornell University.

DANIEL H. COLE, *M. Dale Palmer Professor of Law.* A.B., Occidental College; A.M., University of Chicago; J.D., Lewis and Clark College; J.S.M., J.S.D., Stanford Law School.

JEFFREY O. COOPER, *Associate Professor of Law.* A.B., Harvard University; J.D., University of Pennsylvania.

PAUL N. COX, *Centennial Professor of Law.* B.S., Utah State University; J.D., University of Utah; LL.M., University of Virginia.

ROBIN KUNDIS CRAIG, *Associate Professor of Law.* B.A., Pomona College; M.A., The Johns Hopkins University; Ph.D., University of California; J.D., Lewis and Clark School of Law.

KENNETH D. CREWS, *Associate Dean of the Faculties for Copyright Management and Professor of Law and Library and Information Science.* B.A., Northwestern University; J.D., Washington University; M.L.S., University of California, Los Angeles; Ph.D., University of California, Los Angeles.

JAMES D. DIMITRI, *Clinical Associate Professor of Law.* B.S., Indiana University; J.D., Valparaiso University School of Law.

JENNIFER ANN DROBAC, *Associate Professor of Law*. B.A., M.A., Stanford University; J.D., J.S.D., Stanford Law School.

GEORGE E. EDWARDS, *Associate Professor of Law and Director, Program in International Human Rights Law*. B.A., North Carolina State University; J.D., Harvard Law School.

NICHOLAS GEORGAKOPOULOS, *Professor of Law*. Ptyhion Nomikis, Athens University School of Law; LL.M., S.J.D., Harvard Law School.

JEFFREY W. GROVE, *Associate Dean, Professor of Law, and Director, China Summer Program*. A.B., Juniata College; J.D., George Washington University.

FRANCES WATSON HARDY, *Clinical Associate Professor of Law*. B.S., Ball State University; J.D., Indiana University—Indianapolis.

LAWRENCE A. JEGEN, III, *Thomas F. Sheehan Professor of Tax Law and Policy*. A.B., Beloit College; J.D., M.B.A., The University of Michigan; LL.M., New York University.

HENRY C. KARLSON, *Professor of Law*. A.B., J.D., LL.M., University of Illinois.

ROBERT A. KATZ, *Associate Professor of Law*. A.B., Harvard College; J.D., University of Chicago Law School.

LINDA KELLY-HILL, *Professor of Law*. B.A., J.D., University of Virginia.

ELEANOR D. KINNEY, *Samuel R. Rosen Professor of Law and Co-Director of the Center for Law and Health*. B.A., Duke University; M.A., University of Chicago; J.D., Duke University; M.P.H., University of North Carolina.

ANDREW R. KLEIN, *Professor of Law*. B.A., University of Wisconsin; J.D., Emory University School of Law.

ROBERT E. LANCASTER, *Clinical Associate Professor of Law*. B.A., Millsaps College; J.D., Tulane Law School.

NORMAN LEFSTEIN, *Professor of Law and Dean Emeritus*. LL.B., University of Illinois; LL.M., Georgetown University.

MARIA PABON LOPEZ, *Assistant Professor of Law*. B.A., Princeton University; J.D., University of Pennsylvania Law School.

GERARD N. MAGLIOCCA, *Assistant Professor of Law*. B.A., Stanford University; J.D., Yale Law School.

DEBORAH MCGREGOR, *Lecturer in Law and Assistant Director of Legal Analysis, Research and Communication*. B.A., University of Evansville; J.D., Georgetown University.

SUSANAH M. MEAD, *Associate Dean and Professor of Law*. B.A., Smith College; J.D., Indiana University—Indianapolis.

MARY H. MITCHELL, *Professor of Law*. A.B., Butler University; J.D., Cornell Law School.

JAMES P. NEHF, *Cleon H. Foust Fellow, Professor of Law, and Director, European Law Program*. B.A., Knox College; J.D., University of North Carolina.

DAVID ORENTLICHER, *Samuel R. Rosen Professor of Law and Co-Director of the Center for Law and Health*. A.B., Brandeis University; J.D., M.D., Harvard College.

JOANNE ORR, *Clinical Associate Professor of Law*. B.S., Indiana State University; J.D., California Western.

H. KATHLEEN PATCHEL, *Associate Professor of Law*. A.B., Huntington College; J.D., University of North Carolina; LL.M., Yale University.

FLORENCE WAGMAN ROISMAN, *Paul Beam Fellow and Professor of Law*. B.A., University of Connecticut; LL.B., Harvard Law School.

JOAN M. RUHTENBERG, *Clinical Professor of Law and Director of Legal Analysis, Research and Communication*. B.A., Mississippi University for Women; J.D., Indiana University—Indianapolis.

JOEL M. SCHUMM, *Clinical Assistant Professor of Law*. B.A., Ohio Wesleyan University; M.A., University of Cincinnati; J.D., Indiana University School of Law— Indianapolis.

ANTHONY A. TARR, *Dean and Professor of Law*. B.A., LL.B., University of Natal; LL.M., Cambridge University; Ph.D., University of Canterbury; Ph.D., Cambridge University.

JAMES W. TORKE, *Carl M. Gray Professor of Law*. B.S., J.D., University of Wisconsin.

LAWRENCE P. WILKINS, *Professor of Law*. B.A., The Ohio State University; J.D., Capital University Law School; LL.M., University of Texas.

LLOYD T. WILSON, JR., *Associate Professor of Law*. B.A., Wabash College; M.A., Duke University; J.D., Indiana University—Bloomington.

MARY T. WOLF, *Clinical Professor of Law and Director of Clinical Programs*. B.A., Saint Xavier College; J.D., University of Iowa.

R. GEORGE WRIGHT, *Professor of Law*. A.B., University of Virginia; Ph.D., Indiana University; J.D., Indiana University School of Law—Indianapolis.

Emeriti

EDWARD P. ARCHER, *Professor of Law Emeritus.* B.M.E., Renesselaer Polytechnic Institute; J.D., LL.M., Georgetown University.

AGNES P. BARRETT, *Associate Professor of Law Emerita.* B.S., J.D., Indiana University.

CLYDE HARRISON CROCKETT, *Professor of Law Emeritus.* A.B., J.D., University of Texas; LL.M., University of London (The London School of Economics and Political Science).

DEBRA A. FALENDER, *Professor of Law Emerita.* A.B., Mount Holyoke College; J.D., Indiana University—Indianapolis.

CLEON H. FOUST, *Professor of Law Emeritus.* A.B., Wabash College; J.D., University of Arizona.

DAVID A. FUNK, *Professor of Law Emeritus.* A.B., College of Wooster; J.D., Case Western Reserve University; M.A., The Ohio State University; LL.M., Case Western Reserve University; LL.M., Columbia University.

PAUL J. GALANTI, *Professor of Law Emeritus.* A.B., Bowdoin College; J.D., University of Chicago.

HELEN P. GARFIELD, *Professor of Law Emerita.* B.S.J., Northwestern University; J.D., University of Colorado.

HAROLD GREENBERG, *Professor of Law Emeritus.* A.B., Temple University; J.D., University of Pennsylvania.

WILLIAM F. HARVEY, *Carl M. Gray Professor of Law & Advocacy Emeritus.* A.B., University of Missouri; J.D., LL.M., Georgetown University.

W. WILLIAM HODES, *Professor of Law Emeritus,* A.B., Harvard College; J.D., Rutgers, Newark.

WILLIAM ANDREW KERR, *Professor of Law Emeritus.* A.B., West University; J.D., LL.M., Harvard University; B.S., Duke University.

WILLIAM E. MARSH, *Professor of Law Emeritus.* B.S., J.D., University of Nebraska.

RONALD W. POLSTON, *Professor of Law Emeritus.* B.S., Eastern Illinois University; LL.B., University of Illinois.

KENNETH M. STROUD, *Professor of Law Emeritus.* A.B., J.D., Indiana University—Bloomington.

JAMES PATRICK WHITE, *Professor of Law Emeritus.* A.B., University of Iowa; J.D., LL.M., George Washington University.

Law Library Faculty

JAMES F. BAILEY, III, *Professor and Director of Law Library.* A.B., J.D., M.A.L.S., University of Michigan.

MARY HUDSON, *Reference/Circulation Librarian.* B.A., Ball State; M.L.S., Indiana University.

RICHARD HUMPHREY, *Reference Librarian.* A.A., Brewton-Parker Junior College; B.A., Georgia Southwestern College; M.L.S., University of Kentucky.

WENDELL E. JOHNTING, *Assistant Director for Technical Services.* A.B., Taylor University; M.L.S., Indiana University.

CHRIS E. LONG, *Catalog Librarian.* B.A., Indiana University; M.A., Indiana University; M.L.S., Indiana University.

MAHNAZ K. MOSHFEGH, *Acquisition/Serials Librarian.* B.A., National University of Iran; M.S., Tehran University; M.A., Ball State University; M.L.S., Ph.D., Indiana University.

MIRIAM A. MURPHY, *Associate Director of Law Library.* B.A., Purdue University; J.D., M.L.S., Indiana University—Bloomington.

KIYOSHI OTSU, *Computer System Specialist.* A.A., Parkland College; A.B., M.S., C.A.S., University of Illinois.

Indiana Law Review

Volume 36 **2003** **Number 1**

TRIBUTES

NORMAN LEFSTEIN—SPLENDID DEAN, LEGITIMATE HOOSIER

RANDALL T. SHEPARD*

Few practitioners fully appreciate the challenge of being an effective modern law dean. The dean is surrounded by substantial forces only partially amenable to his or her will. The dean sits as the first among equals in a tenured faculty whose individual interests do not always mesh neatly with the overall goals of the school. Hundreds of students pass by the dean's office each day, many of them ready at the drop of a hat to spring into action proving that they will be superb advocates by doing combat on some matter of student interest. The alumni watch at greater distance, wishing for the school to burnish their own credentials but understanding only in part the trends in legal education. It is no wonder that most law deans last four or five years.

Still, success stories appear before our very eyes. Norman Lefstein's leadership of the school at Indianapolis has been such a story. I record here just a few of the ways in which Dean Lefstein has exceeded himself.

I. REACHING BEYOND THE ACADEMY

American law schools have reformed their education approach at a tremendous pace in the years since the American Bar Association issued the MacCrate Report in 1992.[1] Yet, complaints persist that the academy is too inward looking. Some of these criticisms come from within the academic community itself.[2] Others come from close by, such as a famous complaint by Judge Harry

---

* Chief Justice of Indiana. A.B., 1969, Princeton University; J.D., 1972, Yale Law School; LL.M., 1995, University of Virginia.

1. Robert MacCrate, *Legal Education and Professional Development—An Educational Continuum* (July 1992). The "MacCrate Report," headed by former ABA President Robert MacCrate and presented by the Task Force in Law Schools and the Profession, is a comprehensive examination of multidisciplinary practice in the United States.

2. *See* Donald J. Weidner, *A Dean's Letter to New Law Faculty About Scholarship*, 44 J. LEGAL EDUC. 440 (1994); Donald J. Weidner, *Law School Engagement in Professionalism and Improved Bar Relations*, 72 FLA. B.J. 40 (1998); Donald J. Weidner, *The Crises of Legal Education: A Wake-Up Call for Faculty*, 47 J. LEGAL EDUC. 92 (1997); Donald J. Weidner, *The Florida Supreme Court Commission on Professionalism and the Crises of Legal Education*, 71 FLA. B.J. 64 (1997). Don Weidner served as Dean of Florida State University College of Law from 1991 to 1997, as Interim Dean from 1998 to 2000, and as Dean from 2000 to present.

Edwards, formerly of the law faculty at Michigan: "For some time now, I have been deeply concerned about the growing disjunction between legal education and the legal profession."[3]

To earn a reputation as someone who looks beyond the boundaries of the faculty meeting, a law dean must commit to some heavy lifting: time spent at bar association receptions, travel to relatively small groups of alumni in distant cities, meetings with judges of various stripes, and so on. Norm Lefstein has dedicated a substantial part of his personal energy to building such links during his stewardship of the school.

And he has always given the impression that he thirsts to do more. On integrating the bench and bar into the work of the school, he once said: "We seek to use the practicing bar as mentors to our students, as speakers at a variety of programs, as judges for a wide variety of competitions, and yet I always have the sense that the demand by members of the private bar to be involved in our legal education program exceeds our opportunities to involve them."[4]

Dean Lefstein's efforts along these lines has paid dividends for the school and the profession in a host of ways. His approach has assisted in recruiting both faculty and students, made it easier to integrate adjunct faculty into the teaching of the school, and resulted in broader support for the school's financial needs.[5]

II. COLLABORATION ON LEGAL EDUCATION REFORM

While outreach requires substantial exertion, it is not a particularly risky venture. The same cannot be said of working on alternative methods of legal education. Here, the academy's interests are at a high point, making any decanal

---

3. Harry T. Edwards, *The Growing Disjunction Between Legal Education and the Legal Profession*, 91 MICH. L. REV. 34, 34 (1992):

> For some time now, I have been deeply concerned about the growing disjunction between legal education and the legal profession. I fear that our law schools and law firms are moving in opposite directions. The schools should be training ethical practitioners and producing scholarship that judges, legislators, and practitioners can use. The firms should be ensuring that associates and partners practice law in an ethical manner. But many law schools—especially the so-called "elite" ones—have abandoned their proper place, by emphasizing abstract theory at the expense of practical scholarship and pedagogy. Many law firms have also abandoned *their* place, by pursuing profit above all else. While the schools are moving toward pure theory, the firms are moving toward pure commerce, and the middle ground—ethical practice—has been deserted by both. This disjunction calls into question our status as an honorable profession.

4. John E. Connor & Assocs., 1 Conclave on Legal Education in Indiana 64 (Feb. 28, 1997) (transcript of Conclave on Legal Education in Indiana).

5. Total gifts to the law school during the decade of the 1990s rose from roughly $250,000 a year to about $3 million a year. Dean's Report 2000-2001, Indiana University School of Law—Indianapolis, at 20.

leadership a high-risk endeavor. It has been the characteristic of the Lefstein reign.

When Indiana became one of the early states to stage a "conclave on legal education," Dean Lefstein appeared not to hesitate about co-chairing the project, along with the president-elect of the Indiana State Bar Association, Chic Born.[6] When seventy-five leaders of the bench, the bar, and the academy gathered for nearly two days of discussions about reforming lawyer and law student legal education, it was plain that some of the academics in the hall mostly wanted to get out of town in one piece. It was not Norm's approach. He said:

> [T]here are two final questions that are sure to be before us during this conclave. The first relates to whether or not we are teaching our students what we want them to learn, whether we are effective in doing that. And the second concerns whether there are ways in which legal education should change in order to enhance the confidence of today's law graduates.[7]

One can sort through the leadership ranks in American institutions of all sorts and conclude that we face a shortage of people willing to engage publicly with others about the adequacy of the work their organizations perform. Norm Lefstein has provided us all with a model of forthright, collaborative leadership.

III. LEFSTEIN AND INDIANA'S CRIMINAL LAW PROGRESS

Barbara Allen Babcock will write more fully about Dean Lefstein's contributions to the national effort at improving defense of the indigent in criminal cases.[8] The story of his contributions to Indiana reform, however, should

---

6. *See, e.g.*, William R. Rakes, *Conclaves on Legal Education: Catalyst for Improvement of the Profession*, 72 NOTRE DAME L. REV. 1119, 1124 (1997):

> [T]oday's law school graduates are less prepared for the practice of law than those of two or three decades ago, Lilly agrees the trend toward this theory has created an imbalance in the law schools. He claims that the situation is deteriorating, particularly with regard to the relationship between law faculties and the practicing bar.

Rakes continues,

> I will suggest that beneath this seemingly placid surface lie currents of a major realignment, not between students and faculty, or even between students and practitioners, but rather between the faculties of major law schools and the bench and bar.

Id. (citing Richard Posner, *The Deprofessionalization of Legal Teaching and Scholarship*, 91 MICH. L. REV. 1921 (1993)).

7. Connor & Assocs., *supra* note 4, at 70-71.

8. Dean Lefstein's prominence in this effort is of long standing. *See, e.g.*, Dudley Clendinen, *Budget Ills Crippling Defense of Poor, Lawyers Say*, N.Y. TIMES, Nov. 14, 1982, § 1, at 28 (Whatever the precise measure, it is the apparent inequity, and the erratic nature of indigent defense, that worries Professor Lefstein. "People being arrested are not having their cases challenged in any rational, systematic kind of way," he said, "and the result is, innocent people get

be separately documented.

Larry Landis of the Indiana Public Defender Council led a dogged effort to improve our public defender services, and finally persuaded the General Assembly to adopt a framework for far-reaching reforms in 1989.[9] Governor Evan Bayh appointed Norm as first chair of the new Public Defender Commission of Indiana, and the project has never looked back.

For more than a hundred years, Indiana has held to the ideal that in a decent society a person should not go to trial without a lawyer just because he or she is too poor.[10] It thus befit our heritage that Indiana became the second state in the union to adopt mandatory standards for the appointment of counsel in capital cases (something that still does not exist in federal courts) when the Indiana Supreme Court and the Indiana Public Defender Commission collaborated on court rules and commission standards in 1989.[11] Norm Lefstein captured the enormous effect of these standards in a comprehensive published study.[12]

Perhaps more telling is the effect of the Public Defender Commission's work on the hundreds of thousands of non-capital cases. We have striven to assure indigent criminal defendants that their legal counsel is not doing his or her on-the-job training on them.[13] By today, more than fifty counties are participating in a program to upgrade local public defender services.

This sea change is consistent with what the Indiana Supreme Court said when it first established a right to counsel for indigent defendants. In deciding *Webb v. Baird* more than a century before *Gideon v. Wainwright*,[14] Justice Stuart of our court said:

> It is not to be thought of, in a civilized community, for a moment, that any citizen put in jeopardy of life or liberty, should be debarred of counsel because he was too poor to employ such aid. No Court could be respected, or respect itself, to sit and hear such a trial. The defense of the poor, in such cases, is a duty resting somewhere, which will be at once conceded as essential to the accused, to the Court, and to the public.[15]

Norm Lefstein has played a central role in building upon this foundation by

---

convicted.").

9. The General Assembly created the Indiana Public Defender Commission in 1989 by P.L. 284-1989.

10. Chief Justice Randall T. Shepard, State of the Judiciary, *Counsel, Computers, Compensation, and a Few Words About Dimpled Chads* (Jan. 2001).

11. Norman Lefstein, *Reform of Defense Representation in Capital Cases: The Indiana Experience and its Implications for the Nation*, 29 IND. L. REV. 495, 503-04 (1996).

12. *Id.*

13. Chief Justice Randall T. Shepard, State of the Judiciary, *1995 is Bound To Be A Better Year* (Jan. 1995).

14. 372 U.S. 335 (1963).

15. 6 Ind. 13, 18 (1954). *See also* Randall T. Shepard, *Indiana Law and the Idea of Progress*, 25 IND. L. REV. 943, 947 (1992).

developing the self-respect of these institutions.

IV. THE SCHOOL IS A LARGER PLACE

Everyone I know recognizes that there would not be a fine new law building had it not been for Norm Lefstein. More important yet is the stronger school that rests beneath its roof.

American law schools at the turn of the century were obliged to seek out their own niches in the world of legal education in order to thrive in a market that turned down for most of the decade of the 1990s. Dean Lefstein's vision of the mission of the school has been simultaneously lofty and practical: "To give our students a first-class legal education, to have a faculty that engages in scholarship and research that not only is geared to Indiana but to the nation and the international community as well, and to serve both the academic and legal communities."[16]

Under Dean Lefstein's tutelage, the law school has developed along the lines of the business/environmental adage "Think globally, act locally." It has promoted student involvement at the local level, built national relationships, and strengthened international awareness in the legal community. The dean was instrumental in creating the International and Comparative Law Review and launching Summer Study Abroad programs in places like Lille, France, and Beijing, China. The Program in International Human Rights Law soared as students have represented the school on six continents since its inception. Most recently, the school became home to an LL.M. program in American law for foreign lawyers. The national and international work of the faculty have meant such change in the school's work that it was hardly a surprise that Dean Lefstein's successor came from overseas.

The Center for Law and Health program blossomed to national recognition, the faculty has grown steadily stronger and diverse, and the professional lecture series has attracted notable speakers from U.S. Supreme Court Justices to civil rights activists and dignitaries from other countries. Dean Lefstein has dramatically expanded student opportunities for clinics, internships, and externships.

And, of course, if you build it, they will come. The school's applications have nearly doubled during the Lefstein years. The grade point average of the entering class has risen. Minority enrollment has risen from 0.36% to 22%. And there are more women than men for the first time in history. Moreover, the national and international work of the faculty have expanded the school's reputation so broadly that attracting a successor dean from overseas seemed part of a natural progression

Norm Lefstein has left us an institution of legal education stronger than it has ever been. It is one that makes good of its connections to Indiana, and one that makes noteworthy contributions to the national and international legal scenes.

---

16. Connor & Assocs., *supra* note 4, at 60-61.

V. LONG LIVE NORM!

I learned a new word some weeks back that befits today's occasion: festschrift. A festschrift is a collection of writings from several hands for celebration; especially one of learned essays contributed by students, colleagues and admirers to honor a scholar on a special anniversary.[17] Dean Norman Lefstein is worth celebrating. He was not born in Indiana, but I think he came as soon as he heard about it. He has given us more than we could ever have hoped for. Many thanks, dean, for what you have done and who you are.

17. WEBSTER'S THIRD NEW INTERNATIONAL DICTIONARY 841 (Philip Babcock Grove ed., 1993).

THE LEFSTEIN YEARS
A PRESCRIPTION FOR LEADERSHIP

GERALD L. BEPKO*

A leader has the focus, passion, courage, and wisdom to create an atmosphere that allows others to create.

A leader empowers others, giving them the tools they need and the "room to let them run."

A leader offers compassion and caring to those with whom one's life intersects.

These are the qualities of leadership that promote success, according to Lee Bolman, coauthor of *Leading With Soul: An Uncommon Journey of the Spirit.* They also help to explain the wonderful success of Norman Lefstein's deanship at the Indiana University School of Law—Indianapolis, 1988-2002.

I. FOSTERING CREATIVITY

Norm's particular focus as dean has been obtaining the resources needed to enhance the local, regional, and national reputation of the school. It can almost go without saying, although it shouldn't, that his crowning achievement will be Inlow Hall, the $35 million facility with state-of-the art classroom technologies and research resources. Norm was involved from the very earliest moment of planning, including a crucial meeting in which Norm, then Indiana University (IU) President Tom Ehrlich, and I laid the initial plans for a combined Herron/Law project. Norm was a key factor in securing state funding for the project (to the tune of $21 million), after which he successfully led a capital campaign that raised $14 million in private donations from alumni and other supporters, including the Inlow naming gift. His passion and talent for fundraising also resulted in a steadily increasing annual fund, many privately funded professorships and scholarships, and a new honorary lecture series.

Inlow Hall has not only served the law school community well, it has also become a landmark for the campus. The IUPUI campus has been preparing itself as a great university community for more than thirty years, behind a veil of parking lots. Now, Inlow Hall, the new home of the IU School of Law—Indianapolis, along with the Informatics Complex just to its north, shows a new face to the adjacent downtown area and serves as a new gateway to the campus. The gateway was well conceived by the architects most involved. The Smith Group of Washington, D.C., designed the law school; Robert A. M. Stern of New York designed the Informatics Complex; and Jon Belle, of Buyer, Blinder, and Belle, New York, provided the master planning for this entire

---

* Interim President, Indiana University, and former Chancellor, Indiana University-Purdue University—Indianapolis (IUPUI). The author acknowledges, with gratitude, the research, writing, and editing that Sylvia M. Payne, Special Assistant to the Interim President of IU, contributed to this article.

segment of the campus.

Beyond Inlow Hall, one of Norm's great legacies will be his unwavering and unrelenting support for the creative and scholarly effort of the faculty. Norm conceived and initiated the Program on Law and State Government and supported the continuing development of the Center on Law and Health, which now ranks eighth in the nation[1] and which has become an increasingly important component of IUPUI's status as a home for premiere health-related interdisciplinary research and teaching. He dramatically enhanced the caliber of the faculty through judicious recruitment and retention of the most highly qualified and productive scholars, the creation of endowed professorships, and an emphasis on continuing professional development, thus creating the expectation for, and a working environment conducive to, scholarly productivity. Along with the continuing development of our student body, and the extraordinary accomplishments of our alumni, this advance in scholarship has increased the recognition that the IU School of Law—Indianapolis is one of the nation's truly fine law schools.

The 1996 law school accreditation team praised the significant increase in the quantity of faculty scholarship addressed to national audiences, the high rate of students' passing the Indiana Bar examination, and the expansion of student services (especially career placement services). The external consultants to the Administrative Review Committees that have twice examined Norm's leadership of the school over the past nearly fourteen years praised the leadership and enthusiasm he exhibits in promoting matters of interest to the legal community and to the overall enhancement of the profession. The present strength of the school's clinical programs is directly attributed to him, as is an increased capacity for judicial and governmental internships bearing academic credit.

No review of Norm's extraordinary accomplishments would be complete without mentioning his efforts to diversify the student body. When Norm first began as dean, the school was struggling to recruit minority students. Today it has 139, reflecting an increase in the student diversity ratio from three percent of the school's population in 1988 to nearly twenty percent in 2002. The number of applications overall has increased as well, thus yielding a student body better prepared for success in law school.

It is clear that the faculty, curriculum, and student body are better in every way; that Norm's successor has inherited a stronger school as a result of his tireless and passionate leadership; and that our graduates can take ever greater pride in having earned their degree at the IU School of Law—Indianapolis. The school has always attracted many of our best and brightest college graduates and provided a capstone education for these students right here at the crossroads of Indiana in the state's seat of government and largest population center. The school's value is enhanced by the fact that those who study here are much more likely to end up making their careers and their lifelong contributions right here in Indiana, whether in law practice or the many other roles they play because of the special blend of theoretical and practical education provided by our law

---

1. U.S. NEWS & WORLD REP.

school. All these qualities have matured during the Lefstein years.

II. EMPOWERING OTHERS

No one who was there can ever forget the august presence of U.S. Supreme Court Justice Anthony Kennedy during the September 21, 2001, dedication of Inlow Hall. It was indicative both of the stature Norm has among his peers in the legal profession and of the caliber of graduates who have served our state and nation as lawyers, judges, and lawmakers. On that occasion, Justice Kennedy reminded us: "Our understanding of the law is that it is empowering. It gives you freedom. It gives you the capacity to think, to dream, to hope, to dare."

Norm has this same understanding of the law and has applied it to his leadership of the school. As one faculty member succinctly put it, "He lets us do our jobs." But Norm's notion of empowerment was not as passive as all that. It was more on the model of Ralph Waldo Emerson whose goal as a teacher and mentor was not "to bring men to me, but to themselves." Although the conventional assumption once was that one generation stands on the shoulders of the last in its acquisition of knowledge, we have Emerson to thank for the more interactive model of mentorship and motivational guidance that empowers rather than controls. This is the style of leadership that Norm brought into play to such good effect by giving faculty the tools and freedom they need to be creative and productive in their work.

Early in his deanship, Norm established incentives to attract talented scholars and to seize opportunities when they arose. He expanded the size and influence of the law school's Board of Visitors to increase connections between the needs of the community and the work of the faculty. At the same time, he sought to increase the school's influence in the community by supporting the expansion of clinical programs, with particular emphasis on criminal defense and civil practice. Norm's special commitment to the law school's pro bono program, which was established in 1993, has resulted in students' providing more than 26,000 hours of service to the community, gaining valuable practical experience in the process.

Norm encouraged and supported partnerships with law schools in Beijing, China; Lille, France; La Plata, Argentina; and Queensland, Australia. In addition, he encouraged the creation of the Program in International and Human Rights Law, which has sent student interns to more than thirty-five countries during the past five years. One of the school's most ambitious international ventures, for which Norm advocated ardently, is the new LL.M. Program in American Law for Foreign Lawyers which, beginning this year, will bring international lawyers to Indianapolis for a master's level program while also providing opportunities for our J.D. candidates to interact with attorneys from around the world.

The backdrop for all these achievements has been the broad base of support he has earned and enjoyed because of his excellent rapport with the legal community. The respect Norm has among his peers smoothed the way for the resources to be gathered for many tools of empowerment.

III. SHOWING COMPASSION

At a recent Symposium on Indigent Criminal Defense in Texas, Norm said to the assembled gathering:

> *How we treat the poorest and least powerful members of our society says a whole lot about what kind of society we are. It has always seemed to me, that for lawyers and the justice system, there is nothing more important than what we do in the treatment of our citizens and the protection of their liberties.*

Norm's passion for the betterment of the law school was matched only by his compassion for the indigent and his quiet but steady crusade of advocacy for better programs of criminal defense on their behalf. A nationally recognized expert on legal ethics, Norm has argued eloquently for standards to ensure that effective public defenders are assigned to indigent cases and that continuing legal education be mandatory for those providing representation to defendants unable to afford private attorneys.

As both a former prosecutor and defense attorney, Norm shrewdly cultivated high-profile leadership among members of the bar arguing for a better indigent defense system in Indiana. He was instrumental in achieving reforms in Indiana because he proposed a carrot-and-stick approach to the problem in which the state reimburses cash-strapped counties for forty percent of their cost of providing public defenders if they adhere to standards for providing counsel to poor defendants. To get the reimbursements, counties that appoint lawyers must make sure the lawyers have specific credentials.

In the belief that both the prosecution and defense need to be well represented in the criminal courtroom, Norm, as chair of the Indiana Public Defender Commission from its inception in 1990, worked to secure the independence of the indigent defense function from undue judicial influence in criminal cases and post-conviction death penalty proceedings. Norm has long served as a member of the American Bar Association's standing committee on indigent defense and once served on the national Committee on Criminal Justice in a Free Society with, among others, former U.S. Attorney General Janet Reno, who was then Dade County State's Attorney in Florida. Norm chronicled the result of these experiences in his 1996 publication of *Reform of Defense Representation in Capital Cases: the Indiana Experience and its Implications for the Nation.*[2]

To convey some idea of how delicately Norm balanced his passion for the success of the law school with his compassion for those in need of legal services, but unable to pay for them, I offer a remark that Dean Frank Newton of the Texas Tech Law School made about Norm after he described his efforts to change Indiana's indigent defense system during a symposium on the subject focusing on proposed reforms in Texas:

> *He 's talked to you about the hard part, which is where you come up with*

---

2. 29 IND. L. REV. 495 (1996).

> *the money to change the system. All of us understand that that sometimes adversely affects your relationship politically with counties and state government. The dean lives in the most populous county in Indiana. He has been an advocate for reform in the state of Indiana, yet the legislature recently gave him the money to help build a brand new law school building. So, there really is life after doing the right thing!*

Norm's wry rejoinder to this comment was, "And I had to raise a lot of private money, too."

IV. WHAT TRUE LEADERS GIVE TO US

Lee Bolman concluded his thoughts on the characteristics of a true leader with the story of three stonemasons. When asked what they were doing, the first one said, "I'm cutting stone." The second said, "I'm building a cathedral." The third said, "I'm serving God." A true leader, Bolman said, is the one whose colleagues confidently give the second and third answers to that question because they have been made to feel that their work is meaningful, significant, and enduring.

Norm has created the context in which all those involved in the success of our law school know that their work is meaningful, significant, and enduring. What better legacy could we have? Thank you, Norm, for leading us so effectively and "leading with soul."

LEFSTEIN TO THE DEFENSE

BARBARA ALLEN BABCOCK*

Somewhere along the line, administration has gotten a bad name—synonymous with bureaucracy, red tape and preoccupation with petty concerns. Call someone a great administrator, and your praise is considered faint, or perhaps slightly ironic. But I will risk it because Norman Lefstein *is* a truly great administrator, and the story of how he used his skills to build a struggling little agency into a model of criminal defense is an emblematic one that belongs in any summary of his professional achievements.

It is also a story about the uses of administrative excellence—which like due process of law, does more than merely keep things running along. The story starts in the early 1960s in Washington, D.C. Norman Lefstein, fresh (perhaps fleeing) from several years of civil litigation in Elgin, Illinois, arrived to take a Master's Degree in Trial Advocacy at Georgetown (The Prettyman Program).

Gideon v. Wainwright,[1] assuring a state-paid lawyer to indigent criminal defendants, was still brand new, and the program Norm came to join was one effort to train effective lawyers for the new day coming. Like many other places, the District had no regular public defender agency, but relied instead on lawyers appointed to serve pro bono.

Soon after *Gideon* came down, Congress created a small experimental outfit—we used to call them pilot programs—for providing indigent defense in the Nation's Capital. It was named the Legal Aid Agency, ("the agency" to its first members). Five or six high-spirited young lawyers dedicated themselves to realizing the dream of *Gideon*: of "a vast, diverse country in which every [person] charged with crime will be capably defended, no matter what his economic circumstances, and in which the lawyer representing him will do so proudly, without resentment . . . ."[2]

The agency's problem at its creation (and still) was that the public, and its representatives, do not embrace the dream of *Gideon* for every defendant. Instead, they want their public defenders to represent only the deserving few in court, and to plead the rest guilty. Thus, sooner or later in the life of every public defender agency, its caseload starts to outstrip its resources, and it comes under tremendous pressure to process cases rather than defend them.

This happened to the Legal Aid Agency within a few years of its founding. But the Agency got a second life and grew into the major channel for defense services in the District of Columbia: the Public Defender Service (PDS). Much of the credit goes to Norman Lefstein's administrative brilliance; he became Deputy Director a few months after I took over as Director in 1968. We worked as a team for four years, and then he headed the PDS for three more years.

First on our agenda was to put the agency on a sound statutory and budgetary footing. Norm wrote a model public defender statute, and led the effort to lobby it through a Congress notably unsympathetic to the needs of local citizens. Yet

---

* Judge John Crown Professor of Law, Stanford Law School.

1. 372 U.S. 335 (1963).
2. ANTHONY LEWIS, GIDEON'S TRUMPET 205 (1989).

in his guise of careful administrator rather than crusading defender, Norm talked to them, not about civil rights, but about cost efficiency; not in abstractions but in the details of charts and projections. And Congress responded; the agency grew and prospered. Once more it attracted top legal talent, once more there was a true adversary system at work, and once more poor people had a defender when they faced the state in court.

To keep all this going, Norm needed data; data for his reports, and his regular treks to Congress. We decided that the lawyers must keep records of their work. Now anyone who thinks this was easy does not know public defenders. As I look back on it, herding cats is the right analogy. Defenders consider themselves lawyer-outlaws, iconoclasts, working to preserve precious liberty, instead of fighting over money and keeping records in order to get paid. Freedom from the time clock was the only perquisite of a job short on compensation and prestige. I can still hear the outraged cries, thirty years later. Yet our lawyers kept records—and even submitted to their review—on forms that Norm designed. They did it because they knew his alchemy—how he could turn these facts into a stable future for PDS.

Many of Norm's ideas were novel for the time; today they are the hallmarks of excellence in a defender program. An intensive training program—for instance—to prepare lawyers for the courtroom, for plea bargaining, for counseling, for all the grave responsibilities of defenders. Systematic training using the techniques now familiar from clinical programs, quite new at the time, required considerable resources and planning. Norm Lefstein took it on himself to demonstrate that good training saved time in the long run—on cases reversed for ineffective assistance, on the ability of lawyers to handle a number of cases efficiently.

In the statute he drafted, Norm named the new organization The Public Defender Service. It may have been the first to bear the "Service" title, reflecting the insight that for public defenders, the strictly legal work is only part of the picture. Public defenders need social workers to help in the representation of many clients: to locate treatment and employment opportunities, to counsel on personal issues. Social workers enable the lawyers to present a coherent life picture and plan at sentencing time (an inevitable day for many clients). Norm built an Offender Rehabilitation Program into the PDS statute, along with a provision for trained investigators.

In 1974 the Public Defender Service was named an "Exemplary Project" by the Law Enforcement Assistance Administration of the U.S. Department of Justice; the agency was the only public defender program in the nation to have been recognized in this fashion. Norm still lists this recognition on his official resume; I am here to say it was in large measure his personal accomplishment.

Others are writing about his long service as Dean, but in these years he has not abandoned his old Defender commitments (once a Defender, always a Defender). Again, his successes have the cast of administration: building institutions; writing standards and statutes; guiding and directing programs. Norm Lefstein's resume is a roll call of the important bar and governmental groups working to improve indigent defense services everywhere. For all those who wish to see fully the beauty of the administrative approach to social

injustice, I commend Norm's description of the work of the Indiana Public Defender Commission, a group he continues to chair, in his article, *Reform of Defense Representation in Capital Cases: The Indiana Experience and Its Implications for the Nation.*[3]

I will close with one of my last, and fondest memories of the public defender days that Norm Lefstein and I shared. It was May Day, 1971; anti-war demonstrators threatened to close down the Capital, and marched at rush hour on the various government buildings. Hundreds of people were arrested throughout the morning, and we defense lawyers prepared to represent them. But hours passed, the smell of tear gas faded from the streets, and still no one was brought to court for arraignment. Nor could we find our potential clients in the usual places—the jails, the houses of detention.

Public defenders on motorcycles fanned out over the city, and finally located a thousand people, locked up in the football stadium, without medical, sanitary or other provisions. Night was drawing near and there was a chill in the Spring air. Speedily, Norm drafted up a petition for habeas corpus; without hesitation he called a judge at home to come back to town and hear it. Moonlight was streaming through the courtroom windows as we examined police officers and Justice Department officials and made our case for immediate release. It felt like a great battle over the next few days, as we deployed the defense resources of the city to represent those caught up in the system and unable to help themselves. We were able to do a fine job because we were well-trained and well-organized. And that is due, in great measure, to Norman Lefstein.

Perhaps the reader is wondering about my role as Director of the Agency. I too am an administrator at heart and one who follows the first rule of leadership: get a great deputy. I hired Norm Lefstein. And I did it at a time when the Legal Aid Agency statute set the top salary, that of the Director, at $16,000 per annum. Norm had a young family, and could not live on that amount. "But Norm," I said in persuading him to come, "the statute says nothing about the salary of the Deputy."

---

3. 29 IND. L. REV. 495 (1996).

Indiana Law Review

Volume 36 **2003** **Number 1**

ARTICLES

CALIFORNIA DEATH TRIP

LAWRENCE M. FRIEDMAN*
PAUL W. DAVIES**

There is basically only one way for a person to enter the world; but there are many, many ways to leave it. In some sense, all men and women are born equal, or almost so; and all normal children follow more or less the same trajectory of development. But people die in most unequal ways—some old, some young, some violently, some peacefully, some by accident or disease or otherwise, some in bed, some in hospitals, some alone, some surrounded by family and friends. Death, of course, is the common fate of humanity. No one gets out of here alive.

The title of this article contains a reference to Michael Lesy's odd and disturbing book, *Wisconsin Death Trip*, published in 1973.[1] Lesy reprinted photographs from around the turn of the century made by a photographer in rural Wisconsin named Charles Van Schaick. Interspersed with the photographs are newspaper accounts of suicides, murders, insanity, and other bizarre forms of behavior, from the same general locale. We read, for example, for 1899, that Christ Wold, a farmer, "committed suicide by deliberately blowing off his head with dynamite"; and that "John Pabelowsky, a [sixteen] year old boy of Stevens Point, was made idiotic by the use of tobacco."[2] Lesy's general thesis is this: by the turn of the century, "country towns had become charnel houses and the counties that surrounded them had become places of dry bones."[3] The countryside was, in short, a place of violence and madness; perhaps out of boredom, isolation, and the terrors of social uncertainty. This is one reason, Lesy thinks, for the flight to the cities. Whether Lesy is right or not, the local newspapers he read do record an extraordinary amount of pathological behavior. Much of this behavior ended in sudden or violent death. And sudden or violent death is the realm, par excellence, of the coroner.

There are, as we said, deaths and deaths. Each society has its own way of classifying deaths. Each society considers some kinds of deaths as "normal," and others as unnatural, or even supernatural. In modern society, "normal" death is the death of old, worn out bodies, of people who die in bed or in a hospital. Young people sometimes die, too, and at one time death in childbirth or infancy

---

* Marion Rice Kirkwood Professor of Law, Stanford Law School.
** J.D., 2001, Stanford Law School; Ph.D., 2001, University of California, Berkeley.

1. MICHAEL LESY, WISCONSIN DEATH TRIP (1973).
2. *Id.*
3. *Id.*

or childhood, from cholera, smallpox, diphtheria, and other calamities, were almost if not quite normal; adults too, even in the prime of life, fell victim to such diseases. This became less and less the case as medicine improved its power, and began actually curing people. In any event, there has been and still is a category of deaths that are socially defined as non-normal: murders, suicides, weird accidents, among others. These were grist for the coroner's mill.

The office of the coroner is ancient. It is part of the medieval inheritance of the common law. Shakespeare has a reference to the coroner's inquest ("crowner's quest law") in Hamlet.[4] The American states took over the institution from England, just as they took over the sheriff and the jury system. It seems to have always operated, however, in a kind of obscurity. John G. Lee published, in 1881, a handbook on the work of the coroner in the various states;[5] even then the literature was described as "scanty" and "scattered."

The coroner is still very much a living office in some of the states. It is also, in the opinion of many, something of an anomaly. Massachusetts abolished the position in 1877, and created the post of "medical examiner;" the examiner had to be a medical doctor. New York took this step in 1915. Rhode Island tried having both a medical examiner and a coroner. By the 1990s, most states had either gotten rid of the coroner altogether, and replaced this office with a medical examiner, or with a mixed system of some sort—both a medical examiner and a coroner; or a system in which some counties had coroners, and others had medical examiners.[6] California retains the office of coroner, pure and simple, in many of its counties. But not in all of them. A law of 1969 empowered the Board of Supervisors of the counties to abolish the office "by ordinance" and provide instead for "the office of medical examiner, to be appointed by the said board." The medical examiner was to be a "licensed physician and surgeon duly qualified as a specialist in pathology"; and he would "exercise the power and perform the duties of the coroner."[7] At the beginning of the Twentieth Century, however, the coroner, anomaly or not, was an important official in California's local government. Each county had a coroner. It was, in most counties, an elective office.[8] From 1893 on, the term of office of the coroner was four years.[9]

The literature on the coroner and his work, more than a hundred years after Lee, can still be described as "scanty." Historians have made surprisingly little

---

4. WILLIAM SHAKESPEARE, HAMLET act 5, sc. 1.

5. JOHN G. LEE, HAND-BOOK FOR CORONERS: CONTAINING A DIGEST OF ALL LAWS IN THE THIRTY-EIGHT STATES OF THE UNION, TOGETHER WITH A HISTORICAL RESUME, FROM THE EARLIEST PERIOD TO THE PRESENT TIME (1881).

6. *See* Randy Hanzlick & Debra Combs, *Medical Examiner and Coroner Systems: History and Trends*, 279 JAMA 870 (1998).

7. CAL. GOV'T CODE § 24010 (1969).

8. In Los Angeles County, from 1956 on, the coroner's office was, by law, to be led by a forensic pathologist, whose title was to be "chief medical examiner-coroner." TONY BLANCHE & BRAD SCHREIBER, DEATH IN PARADISE: AN ILLUSTRATED HISTORY OF THE LOS ANGELES COUNTY DEPARTMENT OF CORONER 39 (1998).

9. 1893 Cal. Stat. 367.

use of the files of coroners. Yet these files are of great legal, and social interest. Hence this study. The basic data of this preliminary report consists of the contents of the files of the coroners' inquests in Marin County, California, supplemented by data from two other counties, San Diego and Yolo counties, all from the year 1904. Some data will also be presented from later years in Marin County (1904, 1914, 1924, and 1934). The number of inquests was never great. In Marin, there were twenty-eight inquests in 1904, forty-two in 1914, fifty-four in 1924, and twenty in 1934. Yolo and San Diego were also small counties, with relatively few inquests. By way of contrast, the Coroner of Cook County (Chicago), conducted 3,821 inquests in 1904.[10]

I. MARIN: THE SETTING

Marin County lies just across the Golden Gate from San Francisco. It is linked to San Francisco by a long, narrow, and elegant bridge. The land mass of the county amounts to something more than 500 square miles. Its western border is the fog-bound shore of the Pacific Ocean. The eastern portion is separated from the ocean by a chain of high hills, or low mountains, as you please. Most of the population is concentrated in the towns and cities in the lowlands, along the rim of the north end of San Francisco Bay. Today, the county is booming, and the population is growing fast. The bay is dotted with yachts, house-boats, and pleasure-craft; and new developments crawl up the steep sides of the wooded hills. The coastal towns are bustling centers of the tourist trade; and so too of the cities that rim the Bay, very notably Sausalito, whose shops and restaurants on the water provide views of San Francisco, gleaming in the distance. The population of the county, as of 2000, was 247,289.

Marin at the turn of the century was a much quieter place.[11] There were no bridges linking Marin to San Francisco. The 1890 census counted a mere 13,072 people. Marin at that time had a very high percentage of the foreign-born—men (52.6%) and women (about 30%). Men outnumbered women—69% of the inhabitants were males. Consequently, there were relatively few families. Yolo County was also small (12,684); but mostly native-born. By 1900, Marin's population had risen to 15,702; and the gender imbalance had dropped noticeably—the county was now about 61% male. By 1910, the population had risen to 25,000, and the gender gap had continued to narrow. Yolo County's population hardly rose at all—it was 13,618 in 1900. San Diego County in 1900 had a population of 35,090; about half of these people lived in the city of San Diego itself.

In 1900, both Marin and Yolo counties were mostly rural. Marin had a rural

---

10. Administration of the Office of Coroner of Cook County Illinois: Report Prepared for the Judges of the Circuit Court by the Chicago Bureau of Public Efficiency, at 29 (1911) [hereinafter Cook County Coroner's Report].

11. The source of the information for Marin and Yolo counties is INTER-UNIVERSITY CONSORTIUM FOR POLITICAL AND SOCIAL RESEARCH UNITED STATES HISTORICAL CENSUS BROWSER, *available at* http://fisher.lib.virginia.edu/census.

population of 11,823, and an urban population of 3,879 (if you can call this urban). In Yolo, the rural population was 10,732, the urban population 2,886. In 1900, Marin had eighty manufacturing establishments, Yolo ninety.

No place is "typical," and Marin has its own special character. Many of the deaths in Marin were deaths by drowning; the county is bounded on three sides by water—ocean and bay. It is hard to drown in landlocked Yolo. Marin was also the home of San Quentin prison, an old and famous establishment, and the habitation of many violent men. The prison sits on a spit of land, overlooking the northern end of San Francisco Bay.

The coroner's office, like other offices of the county government, is housed today in the Marin County Civic Center, a stunning building from Frank Lloyd Wright's last years, constructed with great swooping semi-circles on a hilly site on the edge of San Rafael, the county seat. The coroner's office has maintained, virtually intact, all the inquest files from 1852 to the present. From 1904 on, the inquest files usually contain a typed transcript of the proceedings. The Yolo County records contain some typed transcripts, but more often simply a record of the statements of witnesses. The San Diego records (housed in the Research Archives of the San Diego Historical Society) are much skimpier, at least for 1904; they are usually only one page long, and give only the barest essentials of the inquest; transcripts of testimony are rare.

II. CROWNERS' QUEST LAW: THE STATUTES

At the beginning of the Twentieth Century, as we said, laws establishing the office of coroner were still in force in most states. In some of the states, the role of the coroner was quite restricted. In Wisconsin, the coroner was to hold an inquest if the district attorney ordered him to do so, and only if the district attorney had "good reason to believe that murder or manslaughter has been committed."[12] In Utah, inquests were to be held on the deaths of "persons as are supposed to have died by unlawful means;"[13] and in Tennessee, only when there was probable cause to suspect homicide.[14]

The statutes usually set out the basic procedures for coroners to follow. In Illinois, for example, the coroner, "as soon as he knows or is informed that the dead body of any person is found, or lying within the county, supposed to have come to his or her death by violence, casualty, or any undue means," must "repair" to the place where the dead body is located, summon a "jury of six good and lawful men of the neighborhood," and, "upon view of the body . . . inquire

---

12. WIS. STAT. STAT § 4865 (1906).

13. 1907 Utah Laws, tit. 37, § 1221. But in Utah, it was the "justice of the peace" who had the duty to hold inquests.

14. In Tennessee, under TENN. CODE ANN. § 7274 (1896), no inquest was to be held without an "affidavit, in writing . . . signed by two or more reliable persons, averring . . . that there is good reason to believe" that the dead person came to "his, her, or their death by unlawful violence at the hands of some other person or persons."

into the cause and manner of the death."[15] This notion of viewing the body was an essential element of the historic role of coroner's juries; in England, according to Lee, the inquisition was "void," except "*super visum corporis*."[16]

The California statute in force in 1904[17] was somewhat ambiguous on the question of exactly what deaths fell under the coroner's jurisdiction. The statutory trigger read as follows: the coroner steps in when he is "informed" that "a person has been killed, or has committed suicide, or has suddenly died under such circumstances as to afford a reasonable ground to suspect that his death has been occasioned by the act of another by criminal means."[18] We will later discuss exactly what this language means. At any rate, once informed of a death which triggers use the coroner, the coroner was supposed to pick a jury. The minimum number of jurors was six, and the number of jurors varied from case to case, for reasons not very obvious. In Yolo county, six was the normal number; but in San Diego and Marin, there was much more variation. Sometimes there were seven, or even nine or ten jurors. In one case, in 1904 in Marin, a jury of eleven was convened.

The jury was, as we said, required to look at the dead body (many of the inquests were held in funeral parlors), and they could summon and hear witnesses. The jury would pick a foreman, and listen to the testimony of doctors, eye-witnesses, and others. Witnesses were a normal part of the coroner's inquest. In some cases, there were as many as ten witnesses.

It is not clear how the jury was selected—exactly what the process was. No challenges were allowed to the coroner's jury, but a juror who was biased was not supposed to serve on the jury. It does not seem that there was any mechanism for enforcing this rule. What is clear is that many jurors in Marin and Yolo served more than once. Once in a while, the exact same jury would sit on two different inquests (if they were held, for example, on the same day). In 1904, the Coroner of Marin County held multiple inquests on three different occasions, involving seven of the twenty-eight inquests. On one noteworthy day, there were three inquests.[19] Often, one or two jurors would hold over from inquest to inquest. A Chicago report on the Cook County Coroner's office (1911) reported a problem of "professional jurors." The report claimed that some fourteen jurors served on the vast majority of the coroners' inquests in Cook County; and that this was one of the "worst abuses" of the system. The jurors were paid for their labors, and there was a concern that these professional jurors would not exercise independent judgment, but simply do what the coroner wanted.[20]

---

15. 1907 Ill. Laws 213.

16. LEE, *supra* note 5, at 20-21.

17. CAL. PENAL CODE § 1510 (1904).

18. CAL. PENAL CODE § 1510 (1906). The coroner could—and indeed had to—exhume dead bodies if the deaths arose under suspicious circumstances, and the body had already been buried. There were no examples of this in any of the files we examined.

19. This problem—if it was a problem—seemed to get worse over time. In 1934, it was common for the coroner to hold multiple inquests; one day, he held six of them!

20. Cook County Coroner's Report, *supra* note 10, at 8-9, 41-45.

Who were the coroners? The San Diego coroner, Addison Morgan, was in fact a medical doctor.[21] The Marin County coroner, in 1904, F. E. Sawyer, was a funeral director and embalmer who advertised in the local papers.[22] Perhaps he was a doctor as well (he was referred to as "Doctor Sawyer" at least once in the newspapers), but if so, he did not practice. A doctor was nearly always needed at the inquest, and Sawyer nearly always had a doctor available to testify. It appears that undertakers and owners of funeral parlors were, in many states, popular selections or elections as coroners.[23] This continued to be the case. In the early 1950s, in Kentucky, a survey of eighty-two counties found eleven doctors and thirty-one undertakers in the ranks of the coroners (there were also "farmers, farm laborers, taxi drivers, and persons with no occupation"); in Minnesota, however, there were forty-seven doctors, and only twenty-six undertakers (along with a scattering of others—three osteopaths, one dentist, two insurance salesmen, among others).[24]

In Marin and the other counties too, the coroner tended to dominate the proceedings, as far as we can tell. The coroner, or a deputy, sometimes did some investigative work. In one case, concerning Frederick M. Walsh, who drowned in the Bay, the Coroner testified that he tracked down the person from whom the deceased had rented a room in San Francisco, in order to ask him questions[25]; in another case, when an unknown body was found at Angel Island, the coroner put ads in local papers, trying to find out who the man was (with no success).

The inquest was, if the records can be trusted, rather informal, compared to a trial. The jurors were, however, sworn in. Lawyers were not normally present; and the strict rules of evidence were not followed. Jurors could and did ask questions, and some of them seemed to take a more active part in the goings-on than trial jurors would. But the coroner asked most of the questions. He took the leading role in extracting information out of witnesses. Sometimes his statements and questions had a decided slant; and he commented freely on the evidence. Of an Italian man, struck and killed by a train, the coroner remarked, "As far as I can ascertain, he liked his 'vino.'"[26] At an inquest into the death of Mrs. Mattie Jackson, hit by a train at Larkspur, the coroner remarked that he had

---

21. Morgan died on his seventy-eighth birthday; his obituary appeared in the San Diego Union, January 10, 1937.

22. *See, e.g.*, MARIN JOURNAL, Jan. 14, 1904, at 4.

23. Since the coroner has control of the dead body, an undertaker-coroner would be in a terrific position to get the right to do the funeral for the deceased, hence the office of coroner could become a "feeder" for the undertaker's business. At least this was suggested by some observers. *See* Pete Martin, *How Murderers Beat the Law*, SATURDAY EVENING POST, Dec. 10, 1949 (the piece is a general attack on the amateurishness of coroners).

24. NATIONAL MUNICIPAL LEAGUE, CORONERS IN 1953: A SYMPOSIUM OF LEGAL BASES AND ACTUAL PRACTICES (3d ed., May 1955) (unpublished typescript on file with the Stanford Law Library). Funeral directors were also frequent coroners in New Jersey. In Ohio, after 1945, the coroner was required to be a licensed physician.

25. Marin County Coroner's Inquest (MCCI) 752 (Sept. 27, 1904).

26. MCCI 1648 (Dec. 10, 1924).

visited the site with the witness and a representative from the railroad and "I found that his statement, that is, as far as that part of it as to the station was concerned, was absolutely correct."[27] In the case of Michal Grandi, who died in a bakery after eating some meat, the coroner poured cold water on the idea that Grandi had choked to death: "I am positive . . . that he died from chronic alcoholism, and that he was troubled with cirrhosis of the liver and fatty degeneration of the heart. Of course, that could only be determined by an autopsy. Under the circumstances, if you think it unnecessary to have an autopsy, we will render a verdict." A dutiful jury took the hint, and returned a verdict of "acute alcoholism" as the cause of death.[28] At the end of the inquest, the coroner instructed the jury, although these instructions were much less formal than in a regular jury trial. As we have seen, he sometimes almost put words in the mouth of the jurors. In one case, for example, where a woman had died of tuberculosis, the coroner said to the jury: "I think, Gentlemen, it is a clear case of a natural cause of death."[29] However, the jurors were not forced to take the hint; and they did retire outside the presence of the coroner, to deliberate, reach a decision and render a verdict.

The statute, as we saw, was fairly vague on one crucial point—which deaths call for a coroner's inquest? Murder and suicide seem clear enough; but what does "killed" mean? The answer is hardly obvious, and apparently the language gave the coroner considerable leeway. The inquest records show that the coroner interpreted his powers pretty broadly; he conducted an inquest in all sorts of situations where it was not clear whether anybody had been "killed" in the statutory sense. Many inquests were of sudden deaths that, on inquiry, turned out to be from "natural causes." The coroner also investigated quite a few accident cases. Presumably, these were incidents where there was *some* vague chance that a crime had been committed: if not murder, then perhaps recklessness or manslaughter or the like.

By rare good fortune, a reported California case sheds light on the question of the coroner's jurisdiction—and also on the way the coroner actually worked. In 1906, Addison Morgan, the San Diego County coroner—a medical doctor in private practice—sued the county to recover "compensation for his services in some fourteen inquests."[30] The county, apparently, felt it was under no obligation to pay. Its excuse was that the inquests were unnecessary. The coroner described the fourteen cases—in each one there was a sudden death, and the coroner argued that in each one there was at least some hint or possibility of gross neglect, or suicide, or foul play. The court agreed with the coroner, and ordered the fees to be paid. It seems very clear, from the records, that the coroner in Marin County took the same point of view as Addison Morgan. Because of the fee structure, it was clearly in the coroner's interests to stretch a point and look at as many dead bodies as possible.

---

27. MCCI 762 (Dec. 2, 1904).
28. MCCI 1212 (July 8, 1914).
29. MCCI 750 (Aug. 11, 1904) (death of Clara Amelia Ross).
30. The case is *Morgan v. San Diego County*, 86 P. 720 (Cal. Dist. Ct. App. 1906).

This was not exclusively a California problem. The Illinois statute defined the coroner's domain as deaths which came about "by violence, casualty, or any undue means," which is certainly even more ambiguous and opens the door even wider to discretion. Under the Arkansas statute, if the "dead body of any person" was found and the "circumstances of the death" were "unknown," or "if any person die and the circumstances of his death indicate that he has been foully dealt with," the coroner was to become involved. An Arkansas case turned on the same point, more or less, as the San Diego case. A man was sawing wood, "took a fit," fell down and died. The coroner held an inquest, and then sued the county for his fees. In this case, the coroner lost. The court held for the county: "It is not the duty of the Coroner to inquire of sudden deaths, unless there is reasonable ground to believe that they are the result of violence or unnatural means."[31]

Other statutes differed in small or large details from the text of the California law. Some were broader, some were narrower. In Pennsylvania, the coroner came in when the cause of death was "of a suspicious nature and character." In Oregon, there had to be suspicion of criminal means; or of suicide.

III. WHY DID THEY DIE?

The basic question for the coroner's jury was: how did this person meet his or her death. The inquest ends with a verdict. Here is the breakdown of the results (verdicts) of coroners' inquests, in the four sample years in Marin County:

"Natural causes"	27
Suicides	32
Railroad accidents	15
Automobile accidents	12
Drowning	21
"Accidents"	20
Homicides	2
"Other"	15
Total:	144

Of course, we cannot assume that the inquest results were entirely accurate; the jury could make mistakes, or, in some cases, simply lack enough information to come to the right conclusion. Some of the "accidents" could have been suicides; some of the "drowning" entries might have been suicides as well. Many in the "other" category could have been differently classified. But on the whole, we may assume some sort of rough and ready accuracy.

IV. WOMEN AND MEN

What do we learn from the inquest files? Unusual death, at least as far as the coroner was concerned, was a macho business. In the four sample years in Marin County, there were 144 inquests. All except eighteen of the dead bodies were male. This despite the fact that in the entire sample, there were only *two*

---

31. Clark v. Calloway, 52 Ark. 361 (1889).

homicides—a category that would be expected to be heavily male. Scattered data from other places also show, quite uniformly, a preponderance of men. In Baltimore, in the Nineteenth Century, 75% of the inquests were of men.[32] A study of the City of Westminster, England, in the Eighteenth and Nineteenth Centuries found that men outnumbered women two to one, in almost every category of death.[33] Both in Yolo and San Diego, most of the victims were men.[34]

The Marin County data are not discordant with other data. The suicide rates for men were consistently higher than those for women, throughout this period. In 1904, men committed suicide at a rate more than three times that of women; the national rate was 12.2 per 100,000. In 1914, the national rate had risen to 16.1; in 1924, it had dropped to 11.9; in 1934 it was again higher, to 14.9. Most suicides continued to be men, and by more than a three-to-one ratio. For 1934, there were recorded 18,828 suicides in the United States; 14,564 were men, and 4,254 were women.[35]

Men killed themselves under various circumstances and used all sorts of methods. John C. Tait, age forty-three, a native of England, committed suicide, by "self-administered" chloroform, on March 17, 1904. Tait was despondent because he could not find work; he had tried to commit suicide three times before. He wanted to be "buried in a plain wooden box in the common burying ground . . . . I am wholly and solely to blame in this matter";[36] Mathias Enos, a native of the Azores, hanged himself on July 17, 1904, "while suffering from mental trouble";[37] an "unknown white man," who drowned in San Francisco Bay in February, 1914, left a note that said, "Too much rheumatism; not enough money";[38] two men and a woman committed suicide that year "while temporarily insane," two by shooting themselves, one by drowning;[39] Christensen Bungaard, a native of Denmark, thirty-one years of age, was despondent over a girl who rejected him;[40] eighty-two-year-old Rudolph Huber, who was going blind, took strychnine in August 1914.[41] In 1924, Pedro Cano, a twenty-four-year-old

---

32. Suspicious Deaths in Mid-19th Century Baltimore: A Name Index to Coroner Index Reports (Baltimore City Archives) [hereinafter Suspicious Deaths].

33. Maria White Greenwald & Gary 1. Greenwald, *Coroner's Inquests: A Source of Vital Statistics: Westminster, 1761-1866*, J. LEGAL MED. 51, 60 (1983).

34. Coroner's inquests did, however, play a role, at some points of time, and in some places, in investigating the deaths of women who had had illegal abortions. On this point, see LESLIE J. REAGAN, WHEN ABORTION WAS A CRIME: WOMEN, MEDICINE, AND LAW IN THE UNITED STATES, 1867-1973, at 118-29 (1997), reporting Chicago data in the period after the Second World War. There were no examples of abortion deaths in our sample.

35. 2 HISTORICAL STATISTICS OF THE UNITED STATES 414 (1975).

36. MCCI 740 (Mar. 3, 1904).

37. MCCI 748 (July 17, 1904).

38. MCCI 1191 (Feb. 26, 1914).

39. MCCIs 1199, 1200, 1201 (respectively, May 9, 1914, Apr. 23, 1914, Apr. 18, 1914).

40. MCCI 1207 (June 25, 1914).

41. MCCI 1218 (Aug. 11, 1914).

Mexican, an inmate at San Quentin, fractured his skull "by jumping off [third] tier in new prison with suicidal intent";[42] Albert W. Lane, fifty-three, who had "trouble in the head," severed his jugular vein with a razor, in August 1924;[43] Alex M. Olsen, age forty, inhaled gas from a gas stove, and left a note to his wife (who was divorcing him) saying "Now I hope you are satisfied."[44] The only suicide in the 1934 group was Robert Grimes, who threw himself "under an oncoming truck with suicidal intent."[45] These Marin suicides, with three exceptions, were men. Catherine Dubrow, thirty-five, who died on April 25, 1904, was despondent over the death of a child;[46] and Florence Duddy, twenty-two years old, who ingested lysol "with suicidal intent while temporarily insane" and suffering from "melancholia"; her father testified that she was despondent over anemia.[47]

By way of comparison, in San Diego County (1904), there were about thirteen suicides, out of thirty-six coroners' inquests. Possibly one or two others could be included in this category. The inquest papers are often extremely laconic, and in some cases, the cause of death was listed as "unknown." All of the suicides labeled as such were men. Like the men in Marin County, they chose all sorts of ways to kill themselves: Rupert Reisinger took arsenic; Joe Clemens cut his throat with a razor; James Holohan, arrested for drunkenness, hanged himself in jail; W. J. Smith used "illuminating gas"; August Hourteinne took "carbonic acid"; while Filberto Castillo poisoned himself by taking a product called "Rough on Rats." Shooting oneself with a gun was, however, the most popular way out of this earth for these despondent men.[48]

Why is it that men were so much more at risk of killing themselves, or getting themselves killed, than women? The coroners' inquests tended to blame mental illness, "brain trouble," and the like for the suicides—in fact, almost universally. But it is difficult to understand why men should be so much more prone to mental illness than women. Part of the answer to the gender issue might lie in another feature of the inquest records. The men who died were disproportionately immigrants, disproportionately loners, men who lived by themselves, men without obvious family attachments. The 1904 San Diego records included natives of New Brunswick, England, Germany, Norway, the Azores, Switzerland, Wales, Ireland, and China. Eleven of the twenty-eight were foreign born. Most of the Americans were not Californians, but came from somewhere else. Locals tended not to end up in the coroner's files. People with families, homes, connections, jobs, settled routines were less prone to the kinds of sudden or mysterious death that led to the coroner's inquest. And women,

---

42. MCCI 1608 (Feb. 12, 1924).
43. MCCI 1629 (Aug. 30, 1924).
44. MCCI 1652 (Dec. 10, 1924).
45. MCCI 2016 (May 29, 1934).
46. MCCI 743 (May 2, 1904).
47. MCCI 1223 (Oct. 19, 1914).
48. These files are found in the San Diego Historical Society archives, Collection R. 2.69, Box 22.

more than men, had these characteristics. The lonely people, far from home, in boarding-house rooms, were men, not women.

V. ACCIDENTAL DEATH

The information on accidents is, so far, fairly fragmentary. But the issue of accidental death was, apparently, of some importance to the work of the coroner. The coroner's job was to decide whether somebody was responsible (criminally or otherwise) for an accidental death. The goal was to explain, to blame, or exonerate. In one of the 1904 inquests, John Frederick Hansen, who worked on a ship, was struck by a train of the North Shore Railroad. The accident was fatal. The train engineer testified that he saw Hansen on the track, and blew the whistle, but did not have time to stop the train. The verdict: an accident, "and we hereby exonerate the engineer and crew from all blame."[49] In the same year, Alfred Iten, a native of Switzerland, stepped in front of the "gravity car on Mt. Tamalpais Scenic Rr." But the jury said, "we believe his death was due to his own carelessness."[50] Lillien Keefe, nineteen years old, was hit by a train as she walked over a foot crossing. In this case, there was considerable testimony about how the accident happened, and whether it could have been avoided; the general thrust of the questions, however, went toward absolving the engineer of the train, and pinning the blame on Lillien. The verdict: "Being struck by Electric Train at foot crossing . . . and believe no responsibility rests with N.S.R.R.Co for accident."[51] In general, the coroners' juries seemed quite anxious to absolve railroads and other companies from liability. In a rare exception, an inquest in San Diego, in 1904, found that a minister had drowned accidentally, by "falling from a Sale Boat in the Bay of Sandiego." The jury went on to say: "We hereby Recommend that the Harbor Commissioners or those who have Authority to not allow Pleasure Boats or Public Boats carrying Passengers to go out on the Bay or the Ocean without Life presservers."[52] This, of course, did not actually place any *legal* responsibility on anybody in particular. In a Marin case, where an inmate of San Quentin, William Stanley, killed himself with a knife, the coroner's jury recommended that prisoners in "Crazy Alley" not be given

---

49. MCCI 747 (June 21, 1904).

50. MCCI 751 (Aug. 30, 1904).

51. MCCI 756 (June 2, 1904); the very next inquest, into the death of Elmo M. Dempsey, twenty-one, concluded that the cause was "[c]arelessness in attempting to board a train at Larkspur station on the Northshore Electric Rail Road, while the train was in motion." MCCI 757, June 29, 1904. The railroad was exonerated in all four cases of railroad accidents that led to inquests in Marin in 1904.

52. And of course there was the occasional coroner's inquest that *did* find someone culpable; for example, an inquest in Jackson County, Illinois, in 1905, on the death of James Bostic, shot to death by a "night policeman, Fred Jacquot . . . . We find that shooting not justifiable and recommend that Fred Jacquot be held to await the action of the Grand Jury." Coroner's Inquests, Jackson County, Illinois, *available at* http://www.iltrails.org/jackson/coroner1.htm (last visited July 22, 2001).

knives.[53]

The coroner's inquests do thus shed some light on norms of responsibility (or non-responsibility); and they have some relationship to developments in the law of torts. Over time, the meaning of the plain English word "accident" seems to have shifted. In the famous *Farwell* case,[54] for example, the leading case on the fellow servant rule in the United States, Lemuel Shaw uses the word "accident" or "pure accident" to mean an event that was nobody's fault—and for which nobody was really accountable. The United States, particularly in the first half of the Nineteenth Century, could be described as a legal culture of low accountability. All sorts of rules developed, whose thrust was to limit liability for personal injuries—perhaps in order to encourage enterprise; but in any event, sustained by a view that "accidents" simply happened, as bad luck, fate, or the victim's own fault. Over time, a legal culture of high accountability replaced the culture of low accountability. The era of the "liability explosion" (the Twentieth Century) reflects a frame of mine that does not really believe, for the most part, in "accidents," to the same degree and with the same meaning as the earlier period. An "accident" in the Twentieth Century is usually an event that has a cause; and that cause comes to rest on the an organization (or an insurance company) which bears some responsibility for the accident; and will therefore have to pay.

In 1904, this shift was underway but incomplete. For the coroner, "accidental" apparently did not mean mysterious or random or without a cause. But it still implied a lack of legal responsibility. The coroner's work in general assumes that any death, of course, has some sort of cause: death is either "natural," or it calls for some explanation, but the explanation is always in rational, scientific terms.

In the Nineteenth Century, there were many rules of tort liability, but they did not open wide the doors to compensation, in civil cases. Criminal responsibility was at least sometimes a substitute for tort liability in the Nineteenth Century. That is, when the incident was not a pure "accident," there was a tendency to find some *individual* to blame for the occurrence (criminally), or sometimes as an alternative to a civil suit for damages. The very strong trend in the coroners' reports is to blame the victim himself for carelessness, or in any event to excuse a company or corporation.[55] Another example of exoneration, of another sort, is found in a file from Yolo. The dead man is a suspected prowler, shot by a constable. The prowler, who was sixty-nine years old, apparently fired at the constable, who fired back (he said). The coroner's jury found that the constable "was entirely justified in said act."

---

53. MCCI 755 (Oct. 19, 1904).

54. Farwell v. Boston Worcester R.R., 45 Mass. 49 (1842).

55. *See* WILLIAM GRAEBNER, COAL-MINING SAFETY IN THE PROGRESSIVE PERIOD: THE POLITICAL ECONOMY OF REFORM 98 (1976), on the tendency of coroner's juries in West Virginia to exonerate in mine accident cases.

VI. INQUEST FINDINGS AS EVIDENCE

When the coroner's inquest makes a finding of accident, or suicide, or excuses or blames someone for a death, what weight does this verdict have in a court of law? For example, take the case where a coroner's jury brings in a verdict of suicide. What impact does this have in a lawsuit brought by the dead man's family against his insurance company? Many insurance policies provided that the company would not have to pay if the insured killed himself. The formal question was whether the coroner's inquest was "judicial" or "ministerial." If "judicial," the inquest material could be admitted in court. This would not be true if the finding were merely "ministerial." A few cases held the inquest to be "judicial," and hence admissible. *United States Life Insurance v. Volcke*[56] was an Illinois case from 1889; the insured allegedly committed suicide. At least so the coroner's jury found. The court held that the inquest material was admissible as evidence that the dead man killed himself. In 1919, Illinois amended its statute to read that in any negligence case, and in any lawsuit "for the collection of a policy of insurance," the coroner's verdict was not admissible "as evidence to prove or establish any of the facts in controversy." And, indeed, in most states (though not California), the coroner's inquest was *not* acceptable, in cases of this sort.

Aetna Life Insurance Co. v. Milward, a Kentucky case from 1904,[57] was another instance of alleged suicide. Here the court refused to allow inquest evidence to be used in an action against an insurance company. If courts admitted evidence from inquests, the court said, there would be a "race and scramble to secure a favorable coroner's verdict," in order to influence a later tort case, or a claim against an insurance company. Inquests, said the court, are often conducted with "carelessness" and to allow them to be used in a later case would "introduce an element of uncertainty into the practice which would be contrary to public policy, and pernicious in the extreme." This was the prevailing view; it reflects, no doubt, some of the more general suspicion courts had about insurance companies, and their propensity to refuse to pay off claims.[58]

VII. NATURAL CAUSES

In quite a few cases, the coroner's inquest in Marin found that the death was due to natural causes. It is not always clear, in some of these cases, why the coroner was called in at all; we do know (as we mentioned) that it was often to his benefit to investigate, since his income depended on fees. But how often this was a factor is impossible to tell.

Many of these "natural" deaths were, however, rather sudden and therefore

---

56. 129 Ill. 557 (1889).

57. 118 Ky. 716 (1904).

58. *In re L. P. Sly*, 9 Idaho 779 (1904), was a murder case. Sly was accused of murdering one John Hays. After a preliminary examination, he was held without bail on the charge of murder. He filed a writ of habeas corpus, arguing that the proceedings were improper, because no coroner's inquest had ever been held. The Supreme Court of Idaho rejected this argument.

at least vaguely suspicious. In mid-Nineteenth Century Baltimore, 29% of the inquests resulted in a finding of natural causes. The Baltimore coroners investigated, apparently, not just suspicious deaths, but also sudden ones; perhaps another way of putting it, is that a sudden death seemed presumptively suspicious.[59] Typically, an autopsy was held in such cases. In Marin, this happened, for example, in the case of Michael White, an Irishman, whose roommate found him dead in bed. The autopsy doctor decided White had had "hypertrophy of the heart," and a serious kidney problem, caused by drinking; these were what brought on his death. Again, these cases of sudden but natural death were mostly unattached men, who died alone, or in a boarding house. Men with families, attended by doctors, were much less likely to evoke the interest of the coroner, and their deaths would appear "natural" even without an inquest.

VIII. DEATHS FROM MOBILITY

What the coroners' records reveal is the seamy side of American mobility. It was a loose, transient society (for men). It was easy to go off to "seek your fortune"; but lots of men never found this fortune. Just as there were no formal barriers to going up in the world, there were no formal barriers to going all the way down—down as far as it was possible to go. One of the Marin suicides of 1904, John Holtz, a native of Germany, drowned in San Francisco Bay. He was described as a man who once had been wealthy, but had lost his fortune. At the time of his death, he was living in a hotel in San Francisco. He was seventy-four—an old, broken man; and alone.[60] Some men died unmourned and unknown. There were dead bodies that were apparently never identified, like the middle aged man hauled out of the water by a fisherman, in November 1904.[61] Men without family or connections had no way to cushion themselves against disaster, depression, and failure. Even when the death itself turned out not to be abnormal, it was hard to be sure, when a man died alone, without family around him. Alexander Paulsen, a laborer, working on a tunnel, got sick and died: the cause of death was supposed to be "Conjestion [sic] of the lungs." An autopsy was performed; and then "Coroner Sawyer took charge of the remains." Paulsen "was a stranger and no one seems to know anything about him."[62]

There were, during this period, thousands of men (and mostly men), who wandered about in the United States, from place to place. They were looking for work, or a new start, or were simply seized with wanderlust. If they fell toward the bottom rungs of the ladder, they were classified as "tramps" or "hobos" or drifters," and became objects of suspicion and worse.[63] In most states, there were

---

59. *See* Suspicious Deaths, *supra* note 32, at iv.

60. MCCI 760 (Nov. 1, 1904).

61. MCCI 759 (Nov. 1, 1904).

62. THE MARIN JOURNAL, Jan. 21, 1904.

63. *See* PAUL T. RINGENBACH, TRAMPS AND REFORMERS, 1873-1916: THE DISCOVERY OF UNEMPLOYMENT IN NEW YORK (1973); ROGER A. BRUNS, KNIGHTS OF THE ROAD: A HOBO HISTORY (1980).

rather stringent vagrancy statutes; these covered a variety of sins,[64] but were excellent weapons in the police war against tramps. New York passed a specific anti-tramp statute in 1880.[65] Interestingly, the Pennsylvania statute on vagrants and tramps stated specifically that the act was not to apply to any "female."[66] The West in particular was full of "unattached young men" who were looking for work, and "formed a new American underclass," in the late Nineteenth Century.[67]

The California death trip reflects a wider malaise than Lesy found in Wisconsin. Lesy thought the pathologies he found were pathologies of an isolated, rural life. But the same, or worse, pathologies could be found in the cities—and in counties like Marin. Lesy's rural areas, in a way, were pockets of immobility; but the California death trip is much more a tribute to American mobility. Or, if you will, American rootlessness, which is an aspect of the same thing.

Mobility was a central fact of American life—geographic mobility, and also social mobility. From the start, this was a society with its share of risk-takers, entrepreneurs, men (and mostly men) who were trying to climb the greasy pole of success. Sometimes this meant starting a business in one's home town; but often it meant picking up and going somewhere else, to start over, or simply to start. It meant leaving family behind and going to hunt for gold in California. It was a restless society, although it was mostly males who were restless—or who were allowed to be restless. Society encouraged seeking one's fortune. Even the middle class joined in the California gold rush—men who wanted adventure, money, and an escape from the strictures of bourgeois life.[68]

Mobility had an impact on every aspect of society. Among other things, it meant that the population—or a significant part of it—was constantly on the move. Whole communities were made up of strangers; and even in older, settled communities, there were always new people coming in—either from abroad, or from elsewhere in this very big country. Mobility spawned new forms of criminality—forms that depended on a shifting, restless population.[69] Bigamy was one of these crimes—a crime that depended on the ability of men to leave a family behind, and start a new life in some distant community. The strangers in town could include confidence men, sly, cheating men who pretended to be what they were not. Blackmail was another crime that thrived on mobility: it was, in some cases, the crime of threatening to reveal a man's past, in a place where he had started life over again, and thought he had buried that past.

---

64. In the southern states, vagrancy laws were used to control black labor and keep it tied to white landholdings, see, for example, 3 ALA. CODE §§ 6849-50 (1907); the statutes of course did not specifically mention the race issue. *See* William Cohen, *Negro Involuntary Servitude in the South, 1865-1940: A Preliminary Analysis*, 42 J. SOUTHERN HISTORY 31 (1976).

65. RINGENBACH, *supra* note 63, at 23.

66. PA. STAT. ANN. tit. 19290, § 21432.

67. WALTER NUGENT, INTO THE WEST: THE STORY OF ITS PEOPLE 113 (2000).

68. On this, see BRIAN ROBERTS, AMERICAN ALCHEMY: THE CALIFORNIA GOLD RUSH AND MIDDLE-CLASS CULTURE (2000).

69. *See* Lawrence M. Friedman, *Crimes of Mobility*, 43 STAN. L. REV. 637 (1991).

The coroner's bodies represent another aspect of the same mobility. Some at least of the men whose corpses went under the knife, some of the dead bodies that lay in the parlors of undertakers, to be gawked at by the jury—were *victims* of mobility. In many cases, this was literally true: they were killed by railroads, and, later on, automobiles—society's prime instruments of mobility. But in a deeper sense mobility had victimized these men. They were the failures, the losers, the hopeless: men who went off to seek their fortunes, or came to a far-off place to make a start or a fresh start in life; and discovered only sickness, despair, and a lonely death. Their voyage ended in a California death trip.

ENFORCING SETTLEMENTS IN FEDERAL CIVIL ACTIONS

JEFFREY A. PARNESS*
MATTHEW R. WALKER**

INTRODUCTION

Settlements in civil actions in federal district courts may be subject to later judicial enforcement. However, as noted in the 1994 U.S. Supreme Court decision in *Kokkonen v. Guardian Life Insurance Co. of America*, any enforcement "requires its own basis for jurisdiction."[1] Such jurisdiction seemingly can arise under one of two different heads of ancillary jurisdiction in the absence of an "independent basis for federal jurisdiction."[2] One head allows enforcement where the settlement is "in varying respects and degrees, factually interdependent"[3] with a claim that had been presented for adjudication. The other permits enforcement when necessary for the district court "to function successfully, that is, to manage its proceedings, vindicate its authority, and effectuate its decrees."[4]

In *Kokkonen*, there was not a basis for independent jurisdiction and neither head of ancillary jurisdiction supported the enforcement of a settlement that earlier prompted a voluntary dismissal.[5] Any claim for settlement breach had "nothing to do" with any claim earlier presented for resolution, making it neither "necessary nor even particularly efficient that they be adjudicated together."[6] Further, the settlement was not "made part of the order of dismissal";[7] thus, any breach would not be "a violation"[8] of a court order implicating the "court's power to protect its proceedings and vindicate its authority."[9]

Since *Kokkonen*, the lower federal courts have struggled with requests for the exercise of ancillary settlement enforcement jurisdiction. Troubling issues include when and how ancillary enforcement jurisdiction should be retained, when such jurisdiction should later be exercised, and what substantive laws and procedures should be employed in settlement enforcement proceedings. Neither

---

* Professor of Law, Northern Illinois University College of Law. B.A., Colby College; J.D., University of Chicago.

** B.A., Northern Illinois University; J.D., Northern University College of Law.

1. 511 U.S. 375, 378 (1994).
2. *Id.* at 382.
3. *Id.* at 379.
4. *Id.* at 380. Herein, we employ the term "ancillary jurisdiction" as it was used in *Kokkonen*, recognizing that, at times, other terms are used, including pendent, supplemental, residual, derivative, essential, and inherent jurisdiction, as well as jurisdiction of necessity.
5. While the dismissal occurred under FED. R. CIV. P. 41(a)(1)(ii), *id.* at 378, the analysis would have been the same with a dismissal under FED. R. CIV. P. 41(a)(2), *id.* at 381; in both settings, a court order recognizing the settlement was required for any ancillary jurisdiction.
6. *Kokkonen*, 511 U.S. at 380.
7. *Id.* at 381.
8. *Id.*
9. *Id.* at 380.

the Supreme Court in its common law decisions or court rules, nor Congress in statutes, has provided significant guidance. Troubles will likely continue as civil case settlements are being promoted more than ever. The federal district courts recently were expressly directed to facilitate civil settlements and, in order to do so, were authorized to require both party and attorney participation in settlement conferences.[10] After reviewing *Kokkonen* and some contemporary difficulties, we will suggest both lawmaking mechanisms and legal standards for improving settlement enforcement.

I. SETTLEMENT ENFORCEMENT UNDER *KOKKONEN*

Federal district courts are courts of limited subject matter jurisdiction, generally possessing only powers allowed by the federal constitution and authorized by federal statutes.[11] To date, there have been no statutes or court rules governing the retention and exercise of jurisdiction over settlements reached in pending federal civil actions.[12] Given the lack of written laws, some federal courts before 1994 had liberally employed an "inherent powers" doctrine, or similar devices, to enforce settlement agreements reached in civil litigation.[13] Other federal courts were more reticent, leaving most enforcement to the state courts. Some guidance was provided by the U.S. Supreme Court in 1994 in *Kokkonen*. Unfortunately, the ruling in *Kokkonen* addressed only some issues, leaving many questions on settlement enforcement unanswered, and prompting continuing uncertainties and confusion.

The *Kokkonen* case initially involved a dispute over the termination of Matt T. Kokkonen's general agency with Guardian Life Insurance Company.[14] His state court lawsuit was subject to a removal to a federal district court based upon

---

10. FED. R. CIV. P. 16(c). For our thoughts on needed amendments to the rule on settlement conferences in federal civil actions, see Jeffrey A. Parness & Matthew R. Walker, *Thinking Outside the Civil Case Box: Reformulating Pretrial Conference Laws*, 50 KAN. L. REV. 347 (2002).

11. *Kokkonen*, 511 U.S. at 377, 380 (indicating that authorization need not be express, with nonexpress authority sometimes characterized as inherent, ancillary, or essential). There may be small realms of authority beyond congressional control. *See, e.g.*, Eash v. Riggins Trucking Inc., 757 F.2d 557, 562-63 (3d Cir. 1985) (describing "irreducible inherent authority"). *But see* Chambers v. NASCO, Inc., 501 U.S. 32, 48 n.12 (1991) (noting the absence of Supreme Court precedents recognizing such judicial authority).

12. Congress has delegated to the Article III federal courts certain rulemaking responsibilities regarding their own powers. *See, e.g.*, 28 U.S.C. § 2071(a) (2000) (permitting courts to prescribe "rules for the conduct of their business").

13. *See, e.g.*, Lee v. Hunt, 631 F.2d 1171, 1173 (5th Cir. 1980) ("inherent power to enforce"); Kukla v. Nat'l Distillers Prods. Co., 483 F.2d 619, 621 (6th Cir. 1973) ("inherent power").

14. *Kokkonen*, 511 U.S. at 376. Consider: "The complaint, as amended, stated causes of action for wrongful termination, breach of fiduciary duty, interference with prospective business advantage, fraud, breach of lease, wrongful denial of lease, and prayed for damages, including exemplary damages." Petitioner's Brief at *4 n.2, *Kokkonen* (No. 93-263).

diversity jurisdiction where a jury trial was commenced.[15] During trial, the parties reached an oral agreement settling all claims and counterclaims. The key terms of the agreement were recited on the record before the district judge in chambers.[16] "[T]he parties executed a Stipulation and Order of Dismissal with Prejudice"[17] which the district judge signed "under the notation 'It is so ordered.'"[18] The stipulation and order mentioned neither the settlement nor any retention of jurisdiction. When a dispute involving Kokkonen's "obligation to return certain files"[19] under the settlement later arose, Guardian Life moved in the same civil action for enforcement. Kokkonen opposed the motion on the ground that the court lacked subject matter jurisdiction. The district court found it could enforce because it had "an 'inherent power' to do so."[20] The court of appeals affirmed, relying on an "inherent supervisory power."[21]

After noting that the federal courts were "courts of limited jurisdiction,"[22] Justice Scalia, writing for the majority, emphasized that Guardian Life had sought the enforcement of the settlement agreement, not the reopening of the case. He observed that some, but not all, courts of appeals had held that

---

15. *Kokkonen*, 511 U.S. at 376.

16. *Id.* (indicating that "the substance" of the agreement was recited). Guardian Life argued that because of this in camera recitation, the judge "plainly anticipated that any proceeding to enforce the settlement agreement would require an appearance before him and not in state court." Respondent's Brief at *4, *Kokkonen* (No. 93-263). The court of appeals wrote that the "oral agreement . . . was stated in its entirety on the record before the district court in chambers." Kokkonen v. Guardian Life Ins. Co. of Am., No. 93-263, 1993 WL 164884, at *1 (9th Cir. May 18, 1993).

17. *Kokkonen*, 511 U.S. at 376-77.

18. *Id.* at 377.

19. *Id.* Guardian also claimed Kokkonen breached the settlement by communicating to Guardian on behalf of a client who was a Guardian policyholder. Petitioner's Brief at *6 n.8, *Kokkonen* (No. 93-263).

20. *Kokkonen*, 511 U.S. at 377.

21. *Id.*

22. *Id.* Kokkonen framed the issue before the Supreme Court by asking,

> does a federal district court have subject matter jurisdiction to enforce a settlement agreement entered into between the parties when: 1) the case is no longer pending before the court at the time the court issued the order, having been dismissed with prejudice prior to the application for enforcement of the settlement agreement, 2) the settlement agreement has never been incorporated into an order or judgment of the court disposing of the action, 3) the court has not expressly retained jurisdiction over the action, and 4) no other independent grounds for federal court jurisdiction to enforce the agreement exist?

Petitioner's Brief at *i, *Kokkonen* (No. 93-263). Guardian Life framed the issue by asking: "Does a district court have jurisdiction to exercise its discretion to enforce a settlement agreement after dismissal of the case where the settlement was entered into on the record, at trial, with the Court's active participation, and where the Court anticipated its involvement in any enforcement of the agreement?" Respondent's Brief at *i, *Kokkonen* (No. 93-263).

reopening the case in such circumstances was available.[23] In contrast to reopening, Justice Scalia explained that enforcement, "whether through award of damages or decree of specific performance, is more than just a continuation or renewal of a dismissed suit, and hence requires its own basis for jurisdiction."[24] In denying that there was any enforcement power, Justice Scalia cited the absence of an independent basis for subject matter jurisdiction or any ancillary jurisdiction.[25] Yet, Justice Scalia recognized that there were two types of ancillary jurisdiction that might have been available. Ancillary jurisdiction can be exercised "(1) to permit disposition by a single court of claims that are, in varying respects and degrees, factually interdependent . . . and (2) to enable a court to function successfully, that is, to manage its proceedings, vindicate its authority, and effectuate its decrees."[26] Justice Scalia found that any earlier-presented claims and the settlement claim presented by Guardian were not factually interdependent as they had "nothing to do with each other."[27] In the case, he also found that any power to enforce the settlement unaccompanied by a retention of jurisdiction was "quite remote from what courts require in order to perform their functions."[28] He observed that "the only order here was that the suit be dismissed, a disposition that is in no way flouted or imperiled by the alleged breach of the settlement agreement."[29] He noted that

---

23. *Kokkonen*, 511 U.S. at 378 (citing FED. R. CIV. P. 60(b)(6)). The idea of reopening a case was discussed at some length during the oral arguments in *Kokkonen*. Transcript of Oral Arguments, *Kokkonen* (No. 93-263).

24. *Kokkonen*, 511 U.S. at 378. Of course, where a federal civil action, once dismissed, is continued or renewed, there must also be subject matter jurisdiction. Yet, such jurisdiction differs significantly from enforcement jurisdiction in that only with the former is there a return to the claims that prompted the civil action, and thus in effect, a resumption of jurisdiction. Of course, where a state law claim in a federal civil action remains under supplemental jurisdiction after the federal law claims, providing the independent jurisdictional basis is dismissed, there are continuing inquiries into jurisdictional basis. 28 U.S.C. § 1367(c) (2000) (granting courts discretion to decline to continue exercising supplemental jurisdiction).

25. *Kokkonen*, 511 U.S. at 380.

26. *Id.* at 379-80.

27. *Id.* at 380 (concluding "it would neither be necessary nor even particularly efficient that [the claims] be adjudicated together"). Evidently, the claims and counterclaims on which the jury trial was commenced had little or nothing to do with the postjudgment dispute over the return of certain files by Kokkonen. As well, seemingly efficiency would not be promoted by district court settlement enforcement as there was no indication that the district judge was in a unique position to interpret the settlement terms involving the return of the files. *But cf.* Neuberg v. Michael Reese Hosp. Found., 123 F.3d 951, 955 (7th Cir. 1997) (indicating that the judge who presided over the lawsuit was in the "best position to evaluate the settlement agreement"); Scelsa v. City Univ. of New York, 76 F.3d 37, 42 (2d Cir. 1996) ("there are few persons in a better position to understand the meaning of an order of dismissal than the district judge who ordered it").

28. *Kokkonen*, 511 U.S. at 380.

29. *Id.*

> [t]he situation would be quite different if the parties' obligation to comply with the terms of the settlement agreement had been made part of the order of dismissal—either by separate provision (such as a provision 'retaining jurisdiction' over the settlement agreement) or by incorporating the terms of the settlement agreement in the order.[30]

"In that event, a breach of the agreement would be a violation of the order, and ancillary jurisdiction to enforce the agreement would therefore exist."[31] Although the district court "is authorized to embody the settlement contract in its dismissal order (or, what has the same effect, retain jurisdiction over the settlement contract) if the parties agree,"[32] Justice Scalia further wrote that a failure to do so means "enforcement of the settlement agreement is for state courts."[33] The "judge's mere awareness and approval of the terms of the settlement agreement"[34] were insufficient to make those terms a part of the court order, and thus to prompt ancillary jurisdiction.[35]

So, the Supreme Court recognized two ways in which a federal district court could enforce a civil case settlement for a case that had been dismissed.[36] One way involved settlement claims that were factually interdependent with the

---

30. *Id.* at 381. The import of this difference was not said to be reflected in any written federal law. *Cf.* 750 ILL. COMP. STAT. 5/502(d) (2001) (stating that either the terms of a marriage dissolution agreement may be "set forth" in a judgment or that the marriage dissolution case judgment "shall identify the agreement and state that the court has approved its terms," in a setting where such an agreement often is subject to later judicial modification, as where the agreement covers support, custody or visitation of children). This difference has also been deemed important outside the settlement enforcement arena. Smyth v. Rivero, 282 F.3d 268 (4th Cir. 2002) (noting importance to prevailing party status when attorney fee recovery may be available under 42 U.S.C. § 1988 (1994 & Supp. V 1999)). *Compare* Roberson v. Giuliani, 2002 WL 253950 (S.D.N.Y. Feb. 21, 2002) (noting that not all retentions of settlement enforcement jurisdiction prompt prevailing party status under 42 U.S.C. § 1988 (1994 & Supp. V 1999)).

31. *Kokkonen*, 511 U.S. at 381.

32. *Id.* at 381-82.

33. *Id.* at 382.

34. *Id.* at 381.

35. In contrast to federal district courts, when civil actions are settled in the courts of appeal, there is no discretion available to retain jurisdiction over possible settlement breaches. *See, e.g.*, Herrnreiter v. C.H.A., 281 F.3d 634, 637 (7th Cir. 2002) ("a court of appeals lacks factfinding apparatus").

36. Of course, in the absence of a dismissal and a judgment thereon, enforcement could also occur where a pleading was amended to reflect the settlement. *See, e.g.*, Bd. of Managers of the Alexandria Condo. v. Broadway/72nd Assocs., 729 N.Y.S.2d 16 (App. Div. 2001). Yet here too a federal court would need subject matter jurisdiction, often arising under the supplemental jurisdiction statute, 28 U.S.C. § 1367 (2000), because of factual relatedness. *But see* Sadighi v. Daghighfekr, 66 F. Supp. 2d 752, 758 (D.S.C. 1999) (quoting Wilson v. Wilson, 46 F.3d 660, 664 (7th Cir. 1994) ("a district court possesses the inherent or equitable power summarily to enforce an agreement to settle *a case pending before it*") (alteration in original)).

claims presented for court resolution, making adjudication before one trial court "efficient."[37] The other way involved settlement enforcement that promoted successful court functioning. While some found that the analysis in *Kokkonen* led to simple rules,[38] applications of its principles have proven to be difficult. Troubles have already arisen regarding such matters as how to incorporate settlement terms into court orders; how otherwise to retain jurisdiction; whether settlement disputes may prompt the reopening of judgments; and what substantive contract laws and what procedures should apply when federal case settlements are enforced. We find further difficulties in the application of *Kokkonen* which, to date, have gone largely unrecognized. These difficulties include whether there is judicial discretion to refuse party requests that future enforcement jurisdiction be retained, and whether and when any settlement disputes can prompt discretionary refusals to exercise available enforcement jurisdiction.

II. DIFFICULTIES IN SETTLEMENT ENFORCEMENT AFTER KOKKONEN

A. Incorporating Settlement Terms into Court Orders

Under *Kokkonen*, a federal district court may enforce a civil case settlement order after "incorporating the terms of the settlement agreement in the order."[39] Questions have arisen on how settlement terms are properly incorporated. Must all key "terms" be included? If not, which, if any, absent terms are subject to ancillary enforcement jurisdiction? And, what conduct constitutes "incorporation"? The lower courts seem unsure.

The Eighth Circuit has found that a "dismissal order's mere reference to the fact of settlement does not incorporate the settlement agreement."[40] The dismissal order did acknowledge that all matters were settled, but did not otherwise mention the agreement or any of its terms.[41] The appeals court noted that "although *Kokkonen* does not state how a district court may incorporate a settlement agreement in a dismissal order, the case does not suggest the

---

37. *Kokkonen*, 511 U.S. at 380.

38. One commentator suggested that *Kokkonen* "supplies clear guidelines for seeking" supervision of settlement agreements. Charles K. Bloeser, Notes and Comments, *Kokkonen v. Guardian Life: Limiting the Power of Federal District Courts to Enforce Settlement Agreements in Dismissed Cases*, 30 TULSA L.J. 671, 691 (1995). Another said: "For those seeking to ensure federal jurisdiction over agreements settling cases pending in federal court, *Kokkonen* provides a simple answer." Bradley S. Clanton, Note, *Inherent Powers and Settlement Agreements: Limiting Federal Enforcement Jurisdiction*, 15 MISS. C. L. REV. 453, 475 (1995). The petitioner in *Kokkonen* had called "for a 'bright line' rule that will guide district courts in the future." Petitioner's Brief at *17, *Kokkonen* (No. 93-263).

39. *Kokkonen*, 511 U.S. 381.

40. Miener v. Mo. Dep't of Mental Health, 62 F.3d 1126, 1128 (8th Cir. 1995).

41. *Id.* at 1127-28.

agreement must be 'embodied' in the dismissal order."[42] Therefore, the court found that reference to, or even approval of, the settlement agreement was, by itself, insufficient to prompt later enforcement jurisdiction.[43] It did not explain relevant differences between varying nonembodied agreements.

The Ninth Circuit ruled that an order based on a settlement, without more, did not place the agreement within the order.[44] The court stated that the "settlement terms must be part of the dismissal in order for violation of the settlement agreement to amount to a violation of the court's order."[45] Thus, the court concluded that "[w]ithout a violation of the court's order, there is no jurisdiction."[46]

The Sixth Circuit ruled that the "phrase 'pursuant to the terms of the Settlement' fails to incorporate the terms of the Settlement agreement into the order."[47] The lower court had specifically stated: "In the presence of and with the assistance of counsel, the parties placed a settlement agreement on the record before the Hon. Bernard Friedman on October 1, 1991. Pursuant to the terms of the parties' October 1, 1991 settlement agreement, the Court hereby DISMISSES this case."[48]

Some appellate courts have determined that when some, but not all the provisions, of a civil case settlement are placed in a dismissal order, only the incorporated terms are subject to later enforcement proceedings. The Seventh Circuit explained that "[h]aving put some but not all of the terms in the judgment, the district court has identified which it will enforce and which it will not." It further stated that any violation of settlement terms not in a judgment do not "flout the court's order or imperil the court's authority" and thus "do not activate the ancillary jurisdiction of the court."[49] The Tenth Circuit held similarly, stating "[a]lthough the district court specified in its order that it retained jurisdiction, and although it set forth some provisions of the parties' settlement agreement, it did not expressly set forth the provision prohibiting communications to the media."[50] Yet, not all judges may now deny enforcement

---

42. *Id.* at 1128.

43. *Id.*

44. O'Connor v. Colvin, 70 F.3d 530, 532 (9th Cir. 1995).

45. *Id.*

46. *Id.*

47. Caudill v. N. Am. Media Corp., 200 F.3d 914, 917 (6th Cir. 2000). The court cited *In Re Phar-Mor, Inc. Securities Litigation*, 172 F.3d 270, 274 (3d Cir. 1999) (citing Miener v. Mo. Dep't of Mental Health, 62 F.3d 1126, 1128 (8th Cir. 1995) ("The phrase 'pursuant to the terms of the Settlement' fails to incorporate the terms of the Settlement agreement into the order.")). *See also* McAlpin v. Lexington 76 Auto Truck Stop, Inc., 229 F.3d 491 (6th Cir. 2000).

48. *Caudill*, 200 F.3d at 915.

49. Lucille v. City of Chicago, 31 F.3d 546, 548 (7th Cir. 1994).

50. Consumers Gas & Oil, Inc. v. Farmland Ind., 84 F.3d 367, 371 (10th Cir. 1996). Interestingly, the lower court's order of dismissal stated:

> Without affecting the finality of this Judgment in any way, the Court reserves continuing jurisdiction over the implementation and enforcement of the terms of the Stipulation of

of unincorporated settlement terms,[51] especially where breaches of incorporated and unincorporated terms are alleged simultaneously and where all issues are factually interdependent so that their joint resolution promotes efficiency.[52] We favor a bright line test whereby only settlement terms incorporated into court orders (or otherwise referenced particularly) are subject to possible enforcement jurisdiction. Where necessary, efficiency in hearing incorporated and unincorporated pacts together usually can be achieved by a federal court refusal to exercise jurisdiction over the referenced terms, leaving all related matters for a new state court lawsuit.[53]

Under *Kokkonen*, incorporation of settlement terms into a court order is one way to anticipate enforcement jurisdiction. Another way is through a provision retaining jurisdiction over the settlement agreement.[54]

B. Retaining Settlement Enforcement Jurisdiction

Under *Kokkonen*, a federal district court can also enforce if it retains jurisdiction over the settlement agreement.[55] Questions have arisen. Can jurisdiction be retained even though the phrase, 'retaining jurisdiction,' or something like it, is not used? If so, what other terms or actions suffice? At times, are the intentions of the parties and the judge sufficient regardless of the words used? And, can enforcement ever occur after a dismissal where there is no incorporation, no expressly retained jurisdiction, and no subjective intent, but where the exercise of jurisdiction makes sense at the time when enforcement is

---

> Settlement and any issues relating to Subclass membership, notice to Class Members, distributions to Class Members, allocation of expenses among the class, disposition of unclaimed payment amounts, and all other aspects of this action, until all acts agreed to be performed under the Stipulation of Settlement shall have been performed and the final order of dismissal referenced above has become effective or until October 1, 1996, whichever occurs latest.

Id. at 369. It is not clear to us the district judge did not intend to enforce the agreement on media communications, or that its absence is significant given the order's coverage of "all other aspects of this action."

51. *See, e.g.*, Brewer v. Nat'l R.R. Passenger Corp., 649 N.E.2d 1331 (Ill. 1995) (stating the court could enforce a term in the settlement agreement (employee would quit his job) not incorporated into the dismissal order though other terms were included in the order (pursuant to Illinois Code of Civ. Pro. 2-1203, a trial court retains jurisdiction thirty days after entry of judgment)).

52. Of course, in this situation already bootstrapped claims would themselves prompt even more bootstrapping with the unincorporated terms possibly very far removed from the original civil action and perhaps even unknown to the district court until enforcement was sought.

53. Refusals are permitted even when some ancillary enforcement jurisdiction was earlier retained since all ancillary jurisdiction is discretionary. *See* Part III.G, *infra*.

54. Kokkonen v. Guardian Life Ins. Co. of Am., 511 U.S. 375, 381 (1994).

55. *Id. See, e.g.*, Columbus-America Discovery Group v. Atl. Mut. Ins. Co., 203 F.3d 291, 299 (4th Cir. 2000) (stating "court retains jurisdiction to enforce the settlement of the parties").

sought?

The Second Circuit has held that "[o]nce the District Court 'so ordered' the settlement agreement, which included a provision for sealing the case file, it was required to enforce the terms of the agreement,"[56] unless "limited circumstances" permit modification of the "so ordered" stipulation. It reasoned that when a court orders a stipulated and sealed settlement, it accepts certain responsibilities, including a duty to enforce even where there is no court order retaining jurisdiction or incorporating any settlement terms.[57]

In another case, a district judge issued an order stating that any "subsequent order setting forth different terms and conditions relative to the settlement and dismissal of the within action shall supersede the within order."[58] The appellate court stated that "[o]f course, the court may only enter subsequent orders involving the settlement agreement if it has retained jurisdiction."[59] It found that *Kokkonen* "only requires a reasonable indication that the court has retained jurisdiction," as the *Kokkonen* court used the term "such as" when speaking of a separate provision retaining jurisdiction.[60] The court held that the language employed by the district court contemplated a continuing judicial role sufficient to constitute a "separate provision" retaining jurisdiction.[61]

The Eighth Circuit found enforcement jurisdiction was not retained where a dismissal order only stated that the court was "'reserving jurisdiction' to permit any party to reopen the [civil] action."[62] It said that reopening due to a settlement breach was different from enforcing a settlement.[63]

Yet another appeals court ruled that the trial court "need only manifest its intent to retain jurisdiction."[64] The court found this intent in a district court order that declared dismissal was "pursuant to a confidential settlement agreement" and expressly authorized each party to enforce the agreement in the event of breach.[65] The court reasoned "that a district court need not use explicit language or 'any magic form of words.'"[66]

In contrast, a different appeals court held that the mere intent to retain jurisdiction is insufficient.[67] It stated:

At the time the civil case was settled, it is clear that the district court

---

56. Geller v. Branic Int'l Realty Corp., 212 F.3d 734, 737 (2d Cir. 2000).

57. *Id.*

58. Re/Max Int'l, Inc. v. Realty One, Inc., 271 F.3d 633, 645 (6th Cir. 2001).

59. *Id.*

60. *Id.* at 643.

61. *Id.* at 645.

62. Sheng v. Starkey Lab., Inc., 53 F.3d 192, 195 (8th Cir. 1995).

63. *Id.*

64. Schaefer Fan Co. v. J&D Mfg., 265 F. 3d 1282, 1287 (Fed Cir. 2001) (quoting McCall-Bey v. Franzen, 777 F.2d 1178, 1188 (7th Cir. 1985)).

65. *Id.*

66. *Id.*

67. Hagestad v. Tragesser, 49 F.3d 1430 (9th Cir. 1995) (footnote omitted).

intended to retain jurisdiction. It stated at the settlement conference:

> I will act as a czar with regard to the drafting of the settlement papers and the construction of this settlement and the execution of this settlement. And that means that if there is any dispute that is brought to me by counsel, I will decide the matter according to proceedings which I designate in the manner that I designate, and that decision will be final without any opportunity to appeal.

That it believed it had continuing jurisdiction to enforce the agreement is also clear from its order of January 28, 1993:

> As part of the settlement agreement, plaintiff agreed not to provide evidence to prosecute the Oregon State Bar complaint filed against defendant and to take any and all reasonable actions to prevent that matter from proceeding. The parties also agreed that the terms and conditions of the settlement agreement were to remain confidential and not disclosed to anyone. The parties further agreed that all questions relating to their rights and duties under the agreement would be determined exclusively by the undersigned.

> It is equally clear, however, that the district court did not retain jurisdiction over the settlement. As noted, the Dismissal neither expressly reserves jurisdiction nor incorporates the terms of the settlement agreement.[68]

This holding was later reaffirmed when the same court held that "even a district court's expressed intention to retain jurisdiction is insufficient to confer jurisdiction if that intention is not expressed in the order of dismissal."[69]

In the absence of incorporation, jurisdiction retention, or intent, judicial enforcement of settlements still seems appropriate in certain settings. Parties to a federal civil action ending in a judgment upon a settlement are unable to return to the district court with an agreement indicating a new-found intent that jurisdiction over an earlier settlement be retained.[70] Yet, so long as a federal civil

---

68. *Id.* at 1433.

69. O'Connor v. Calvin, 70 F.3d 530, 532 (9th Cir. 1995).

70. *See, e.g.*, Lane v. Birnbaum, 910 F. Supp. 123 (S.D.N.Y. 1995). The court stated:

> In this case, the Order of Dismissal preceded the Stipulation by almost two months. It is therefore apparent that compliance with the agreement was not an operative part of the dismissal. That the parties subsequently felt the need to have the terms of their agreement embodied in a stipulation on file with the Court, cannot serve to vest the Court with jurisdiction over the agreement. . . . Clearly, the Court's dismissal of the action was in no way conditioned upon the parties' compliance with the terms of the agreement. Nor did the Court retain jurisdiction over the parties' agreement. Therefore, enforcement of the settlement agreement is a matter of contract between the parties, for

action remains open because there is no final judgment, a district court seemingly may enforce a settlement therein even though the judge never earlier considered enforcement.[71] Thus, in dismissing a civil action upon a settlement, a trial judge may reserve rendering a judgment as by granting a conditional dismissal, thereby allowing a party to return to court for any reason, including settlement enforcement, before a final judgment is entered.[72]

C. Discretionary Refusals of Later Settlement Enforcement Jurisdiction

Where any later settlement enforcement would not have "its own basis for jurisdiction,"[73] thus requiring some form of ancillary power, can a federal district judge refuse to incorporate the settlement terms into a court order or otherwise to retain enforcement jurisdiction though requested by all parties? The Supreme Court in *Kokkonen* said that with any dismissal of a pending civil action based on a settlement,[74] potential enforcement is "in the court's discretion."[75] This comports with the longstanding principle that ancillary jurisdiction is discretionary. What factors should guide such exercises of discretion?

One appeals court has urged caution when a federal district judge decides

---

the state courts to address.

Id. at 128 (footnote omitted).

71. *See, e.g.*, Sadighi v. Daghighfeker, 66 F. Supp. 2d 752 (D.S.C. 1999). The court stated:

> [A]fter the court was informed that settlement had been reached, there was a delay when no formal settlement documents were executed and no order of dismissal was issued. Consequently, when Defendants decided that the settlement agreement reached earlier was no longer to their satisfaction, the case was still on [the] court's active docket . . . . In short, nothing had been done to divest [the] court of jurisdiction.

Id. at 758.

72. *See, e.g.*, Bell v. Schexnayder, 36 F. 3d 447, 450 n.2 (5th Cir. 1994) (stating that *Kokkonen* is "distinguishable from our case, since here the district court's order of dismissal expressly provided that the parties could, within 60 days, move to reopen the case to enforce the settlement. Defendants so moved within the 60 days of the dismissal order."). Similar trial court initiatives can be addressed in court rules. *See, e.g.*, Form 7-345 of Florida Small Claims Rules ("Stipulation for Installment Settlement, Order Approving Stipulation, and Dismissal," under which proceedings are stayed by agreement while settlement monies are paid over time, with an expressly recognized enforcement power). Yet, conditional dismissal orders, without judgments, may permit later settlement enforcement proceedings. *See, e.g.*, Pratt v. Philbrook, 38 F. Supp. 2d 63, 66 (D. Mass. 1999) (stating conditional dismissal grounded on settlement where parties have sixty days to return "to reopen the action *if settlement is not consummated by the parties*"); *see also* Pratt v. Philbrook, 109 F.3d 18, 21 n.5 (1st Cir. 1997) (stating that the sixty-day procedure developed as a mechanism to close cases "while retaining jurisdiction to enforce a settlement for a period of time after closure is announced").

73. Kokkonen v. Guardian Life Ins. Co. of Am., 511 U.S. 375, 378 (1994).

74. *See, e.g.*, FED. R. CIV. P. 41(a)(1)(ii) ("stipulation of dismissal signed by all parties") and FED. R. CIV. P. 41(a)(2) (dismissal "upon order of the court").

75. *Kokkonen*, 511 U.S. at 381.

whether to enter a consent decree. The Fifth Circuit stated that "[t]he court, however, must not merely sign on the line provided by the parties."[76] The court opined that though a proposed decree has the consent of the parties, the judge should not give perfunctory approval because the court's duty is akin, but not identical to its responsibility in approving settlements of class actions, stockholders' derivative suits, and proposed compromises of claims in bankruptcy."[77] The appeals court declared that the trial court must ascertain whether the settlement is fair, adequate, and reasonable.[78] Where a proposed consent decree, "by virtue of its injunctive provisions, reaches into the future and has continuing effect," the terms require careful scrutiny,[79] presumedly because a trial court is "a judicial body, not a recorder of contracts."[80]

Another appeals court ruled a trial court must "ensure that its orders are fair and lawful," meaning that an agreement that is made part of an order necessarily has judicial imprimatur and contemplates judicial "oversight."[81]

For settlements that are not incorporated into court orders, but over which enforcement jurisdiction may be retained, does discretion operate differently? If so, should trial judges scrutinize such terms more or less carefully? While these settlements are not consent decrees, they are also not wholly private agreements.[82] For us, it seems that in all settings district judges should exercise at least some discretion before agreeing to enforce a civil case settlement agreement if a dispute arises later.[83] Thus, where enforcement jurisdiction is retained but the settlement is not formally filed (as a record available to the public),[84] a copy of the settlement should not only be provided to the court, but the court should also determine it is an appropriate subject for possible court enforcement and oversight, though its terms normally do not need to receive full judicial approval.[85]

---

76. United States v. City of Miami, 664 F.2d 435, 440 (5th Cir. 1981) (footnotes omitted).

77. *Id.* at 440-41.

78. *Id.* at 441 n.13 (requiring further that the agreement must also have the valid consent of the concerned parties and be "appropriate under the particular facts," meaning "a reasonable factual and legal determination based on the facts of record").

79. *Id.* at 441 (stating further that the agreement cannot violate the "Constitution, statute, or jurisprudence").

80. Ho v. Martin Marietta Corp., 845 F.2d 545, 548 n.4 (5th Cir. 1988).

81. Smyth v. Rivero, 282 F.3d 268, 282 (4th Cir. 2002).

82. *See, e.g., id.* at 280 ("a private settlement, although it may resolve a dispute before a court, ordinarily does not receive the approval of the court").

83. For example, enforcement jurisdiction should not be retained where later disputes inevitably would involve novel or complex issues of state law, or where there are "compelling reasons for declining jurisdiction. 28 U.S.C. § 1367(c)(1) & (4) (2000).

84. Jessup v. Luther, 277 F.3d 926 (7th Cir. 2002) (intervenor granted access to civil rights settlement agreement that had been submitted for court "approval" and maintained under seal in court's file even though jurisdiction to enforce it was not retained).

85. *See, e.g.*, Roberson v. Giuliani, 2002 WL 253950, at *2 (S.D.N.Y. Feb. 21, 2002) (contract "provided" to court, but not filed or subject to "so ordered" judgment). Certainly, judges

D. Reopening Federal Civil Actions

Under *Kokkonen*, a district court is enabled, in ruling on a Rule 60(b) motion to set aside a judgment, to influence, if not exercise jurisdiction over, a breached settlement that had previously ended a civil action.[86] If a breach of a settlement can prompt post judgment relief overturning the settlement by reinstating the claims, even though the settlement was never incorporated into the judgment and enforcement jurisdiction was not otherwise retained, in most instances a new settlement will simply follow.[87]

Prior to *Kokkonen*, the appellate courts were split on whether such a settlement breach provided sufficient reason to grant a motion for judgment modification.[88] In *Kokkonen*, the court did not address the issue, finding "that

---

should never agree to enforce illegal or procedurally unconscionable settlement agreements. And at times, in order to ensure fairness to certain parties, as with class actions and claims by minors, judicial approval of the substance of settlements is required.

86. Federal Rule of Civil Procedure 60 is entitled "Relief from Judgment or Order"and reads in part:

> (b) Mistakes; Inadvertence; Excusable Neglect; Newly Discovered Evidence; Fraud, Etc. On motion and upon such terms as are just, the court may relieve a party or a party's legal representative from a final judgment, order, or proceeding for the following reasons: (1) mistake, inadvertence, surprise, or excusable neglect; (2) newly discovered evidence which by due diligence could not have been discovered in time to move for a new trial under Rule 59(b); (3) fraud (whether heretofore denominated intrinsic or extrinsic), misrepresentation, or other misconduct of an adverse party; (4) the judgment is void; (5) the judgment has been satisfied, released, or discharged, or a prior judgment upon which it is based has been reversed or otherwise vacated, or it is no longer equitable that the judgment should have prospective application; or (6) any other reason justifying relief from the operation of the judgment.

87. We think such reopened cases have final settlement rates at least comparable to those for other civil cases. In any event, it seems clear that most reopened cases will eventually settle, if they do not otherwise end without trial.

88. *Compare* Fairfax Countywide Citizens v. County of Fairfax, 571 F.2d 1299, 1302-03 (4th Cir. 1978) (footnote omitted) (holding that "upon repudiation of a settlement agreement which had terminated litigation pending before it, a district court has the authority under Rule 60(b)(6) to vacate its prior dismissal order and restore the case to its docket"), *with* Sawka v. Healtheast Inc., 989 F.2d 138, 140 (3d Cir. 1993) ("Assuming arguendo that Healtheast breached the terms of the settlement agreement, that is no reason to set the judgment of dismissal aside, although it may give rise to a cause of action to enforce the agreement. Relief under Rule 60(b)(6) may only be granted under extraordinary circumstances where, without such relief, an extreme and unexpected hardship would occur.") *See also* Keeling v. Sheet Metal Workers Int'l Ass'n, 937 F.2d 408, 410 (9th Cir. 1991) ("Repudiation of a settlement agreement that terminated litigation pending before a court constitutes an extraordinary circumstance, and it justifies vacating the court's prior dismissal order."); Harman v. Pauley, 678 F.2d 479, 481-82 (4th Cir. 1982) (in this case "interests of justice do not require vacation of dismissal order"); Aro Corp. v. Allied Witan Co., 531 F.2d 1368, 1371

what respondent seeks in this case is enforcement of the settlement agreement, and not merely reopening of the dismissed suit by reason of breach of the agreement that was the basis for dismissal."[89] The court noted that settlement enforcement, "whether through award of damages or decree of specific performance," was different[90] because it was "more than just a continuation or renewal of the dismissed suit"[91] and thus required its own basis for jurisdiction.[92]

After *Kokkonen*, the Sixth Circuit foreclosed a Rule 60(b) motion founded on an alleged settlement breach. The court said that the rule could not support enforcement of a settlement agreement not expressly incorporated in a court order because relief from a final judgment was an extraordinary remedy available only in exceptional circumstances.[93] The request for a contempt finding was deemed "clearly 'more than just a continuation or renewal of the dismissed suit'" and any use of the judgment modification rule would "create an exception to the holding in *Kokkonen* that would swallow the rule."[94]

The Seventh Circuit has held that "[n]othing in *Kokkonen* purports to change the stringent standards that govern the availability of relief under Rule 60(b)(6),"[95] so that a movant could not, in the guise of attempting to set aside an order, seek judicial interpretation of a settlement that was not incorporated in a court order and over which there was no retained jurisdiction.[96]

However, like the pre-*Kokkonen* split, there may now be a post-*Kokkonen* split. One federal district court, after referencing *Kokkonen*, found "that federal courts are empowered to reopen suits dismissed by reason of breach of a settlement agreement by virtue of Rule 60(b)(6)."[97] Another court allowed a

---

(6th Cir. 1976) (court had full power to vacate its order of dismissal when one party "attempted repudiation of the agreement on which the dismissal rested").

89. Kokkonen v. Guardian Life Ins. Co. of Am., 511 U.S. 375, 378 (1994).

90. *Id.* Of course, there must also be some jurisdictional basis for a Rule 60(b) motion, though such a basis was not discussed in *Kokkonen*. Authority over judgment modification motions is rarely questioned on jurisdictional grounds.

91. *Id.*

92. *Id.* Judgment modification was discussed during the oral arguments in *Kokkonen*. *See* Transcript of Oral Arguments, *Kokkonen* (No. 93-263).

> How about any other 60(b)(6), the catch all, and the judge saying well, it sounds like a pretty good 60(b) motion to me; I was listening to these two people debate what their settlement was going to be, and they made certain representations, and one of them is trying to get out of it. So I think that fits the 60(b)(6) catchall. It justifies relief to tell me one thing and the [sic] go do another thing.

Id.

93. McAlpin v. Lexington 76 Auto Truck Stop, 229 F.3d 491, 502-03 (6th Cir. 2000).

94. *Id.* at 503.

95. Neuberg v. Michael Reese Hosp. Found., 123 F.3d 951, 955 (7th Cir. 1997).

96. *Id.*

97. Trade Arbed Inc. v. African Express 941 F. Supp. 68, 70 (E.D. La. 1996) (emphasis omitted). *See also* Rovira v. Fairmont Hotel, 1997 WL 707115, at *2 (E.D. La. Nov. 12, 1997) ("In *Kokkonen*, the Supreme Court ruled that federal courts do not have the power to enforce settlement

Rule 60 motion in a more unusual setting; the case involved a settlement that had been reached between the parties before the court entered a judgment based upon a pending motion. The judge explained that as the "parties' settlement agreement preceded the entry of judgment [upon the grant of the motion], by the clerk of this court the plaintiff is entitled to postjudgment relief pursuant to Fed. R. Civ. P. 60(b)(1) . . . on the grounds of mistake."[98] The court further explained "[i]t would be this court's mistake of fact, i.e., that the parties had not settled the claims at bar before entry of judgment . . . that justifies relief."[99] Instead of reopening the case, the district judge withdrew its ruling and gave the defendant "thirty-five (35) days . . . to comply with the terms of the settlement agreement."[100] The court stated that if the defendant failed to comply, "the plaintiff may return . . . for whatever relief is appropriate."[101]

E. Choosing the Applicable Contract Laws

When *Kokkonen* permits settlement enforcement, questions have arisen about which contract laws apply. The Seventh Circuit recently ruled that "[t]he uncertainty . . . over whether state or federal law would govern a suit to enforce a settlement of a federal suit, has been dispelled; it is state law."[102] This ruling applies to settlements involving both federal and state law claims.[103] Yet, most rules have exceptions and therein lies the rub. Helpful guidelines on any exceptions to state law applicability are hard to find. A second appeals court has simply declared that state contract law operates "unless it presents a significant conflict with federal policy,"[104] with such conflicts "few and restricted."[105] Another appeals court was more specific, holding that local law applies unless the settlement is sought to be "enforced against the United States" or there was

---

agreements that produce stipulations of dismissal. . . . This ruling, however, does not prevent federal courts from reopening dismissed suits when the interests of justice justify such relief."); Hernandez v. Compania Transatlantica, 1998 WL 241530, at *2 (E.D. La. May 7, 1998) ("Federal Rule of Civil Procedure 60(b)(6) empowers a federal district court to reopen a dismissed suit due to a party's breach of a settlement agreement.").

98. Davis v. Magnolia Lady Inc., 178 F.R.D. 473, 474 (N.D. Miss. 1998).

99. *Id.* at 474-75 (also relying on Rule 60(b)(6)) (emphasis omitted).

100. *Id.* at 476.

101. *Id.*

102. Lynch v. Samatamason, 279 F.3d 487, 490 (7th Cir. 2002).

103. *See, e.g.*, United States v. McCall, 235 F.3d 1211, 1215 (10th Cir. 2000) (federal question claim involving issue of whether a settlement offer extended by the Assistant U.S. Attorney was accepted by appellee); Carr v. Runyan, 89 F.3d 327, 331 (7th Cir. 1996) (diversity claim where issue on appeal was whether daughter had the authority to bind mother to settlement agreement reached in mediation).

104. Ciramella v. Reader's Digest Ass'n, Inc., 131 F.3d 320, 323 (2d Cir. 1997) (citing Atherton v. FDIC, 117 S. Ct. 666, 670 (1977)).

105. *Id.* (quoting O'Melveny & Myers v. FDIC, 512 U.S. 79, 87 (1994).

"a statute conferring lawmaking power on federal courts."[106]

The exceptional conditions under which federal laws apply to settlements of federal civil actions are difficult to discern from Supreme Court precedents. In one case, federal decisional contract law on the validity of a written prelawsuit release of a federal statutory claim, allegedly procured by fraud, was applied to the settlement of a case filed in a state court because otherwise "federal rights . . . could be defeated," because settlements of claims under that federal law "play an important part" in the "administration" of the relevant federal act, and because if "federal law controls," there would be "uniform application throughout the country essential to effectuate" the purposes underlying the federal statutory right to sue.[107] And, in another case involving a different federal statutory claim presented in a state tribunal, the high court simply said that "waiver" of the "right to sue" was governed by federal law because "the policies underlying [the federal statute may] in some circumstances render that waiver unenforceable."[108]

Based on such precedents, there are times when federal district courts should employ federal contract law principles in reading federal case settlement agreements. One district court nicely summarized the relevant factors.[109] They include: 1) whether Congress has expressed a policy of encouraging voluntary settlement of the relevant federal statutory claims; 2) whether "the Supreme Court has already articulated certain prerequisites to the validity of settlement agreement" of any relevant federal claims; 3) whether any settled federal claims are within exclusive federal court subject matter jurisdiction; 4) whether state laws in the relevant area of law are preempted "through a comprehensive statutory scheme"; 5) whether there is an expressed federal governmental interest "in remedying unequal bargaining power" between the settling parties; 6) whether the United States is a party to the settlements; and 7) whether Congress empowered the federal courts "to create governing rules of law."[110]

When state contract laws are employed to sustain and interpret settlement agreements reached in federal civil actions, difficulties can arise because the sources of state law extend far beyond the "substantive" matters demanded by the Erie doctrine. Specifically, some state written civil procedure laws, seemingly operative only in the state trial courts, are used in the federal district courts. For example, federal courts have utilized a Texas Rule of Civil Procedure which

---

106. Makins v. District of Columbia, 277 F.3d 544, 547-48 (D.C. Cir. 2002).

107. Dice v. Akron, Canton & Youngstown R. Co., 342 U.S. 359, 361-62 (1952) (claim under the Federal Employers' Liability Act). The decision seemingly was not followed in *Good v. Pennsylvania Railroad Co.*, 384 F.2d 989 (3d Cir. 1967) (state law governs lawyer's authority to settle client's FELA case) and *Pulcinello v. Consolidated Rail Corp.*, 784 A.2d 122 (Pa. Super. Ct. 2001) (FELA case settlement governed by state law on validity of oral agreements).

108. Town of Newton v. Rumery, 480 U.S. 386, 392 (1982) (civil rights claim under 42 U.S.C. § 1983). The decision was criticized in Michael E. Solimine, *Enforcement and Interpretation of Settlements of Federal Civil Rights Actions*, 19 RUTGERS L.J. 295 (1988).

109. Sears, Roebuck & Co. v. Sears Realty Co., Inc., 932 F. Supp. 392 (N.D.N.Y. 1996).

110. *Id.* at 398-401.

states "no agreement between attorneys or parties touching any suit pending will be enforced unless it be in writing, signed and filed with the papers as part of the record, or unless it be made in open court and entered of record."[111] And at times, but not always, federal courts employ state professional conduct and civil procedure law standards to determine the authority of a person other than the party to settle pending civil actions on behalf of that party.[112]

F. Choosing the Applicable Procedures

When a district court exercises jurisdiction over an alleged breach of a civil case settlement there are a variety of procedures that may be used. Possible procedures appear in the Federal Rules of Civil Procedure as well as in common law decisions and statutes.[113] Some, but not all, procedures are geared toward enforcement and remedies on behalf of the party harmed by the settlement breach.

For some settlement breaches, the court may proceed in contempt.[114] There are two forms of contempt, civil and criminal,[115] and either form may be direct or indirect. The major goals of criminal contempt are less connected to enforcement, as they chiefly involve punishment and vindication.[116] On the civil

---

111. *In re* Omni, 60 F.3d 230, 232 (5th Cir. 1995) (quoting TEX. R. CIV. P. 11). The Texas rules are said to "govern the procedure in the justice, county, and district courts of the State of Texas in all actions of a civil nature, with such exceptions as may be hereinafter stated." TEX. R. CIV. P. 2. A similar New York provision, CPLR § 2014, has prompted "disagreement" over its applicability to federal civil actions in the Court of Appeals for the Second Circuit. Turk v. Chase Manhattan Bank USA, NA, No. 00CIV1573CMGAY, 2001 WL 736814, at *2 n.1 (S.D.N.Y. June 11, 2001).

112. *Compare* United States v. Int'l Bhd. of Teamsters, 986 F.2d 15, 20 (2d Cir. 1993) (federal precedent regarding attorney settlement authority used); Reo v. U.S. Postal Serv., 98 F.3d 73, 77 (3d Cir. 1996) (under Federal Tort Claims Act, state law used to determine settlement authority of representative of a child); Neilson v. Colgate-Palmolive Co., 993 F. Supp. 225, 226-27 (S.D.N.Y. 1998) (pursuant to local federal rule, court dispenses with certain state law requirements governing Guardian Ad Litem's power to settle a civil case on behalf of adult incompetent to pursue her own claims as technical compliance with state law would prompt "extended and prejudicial delay").

113. *See* 18 U.S.C. § 401 (2000) (criminal contempt); FED. R. CIV. P. 65 (injunctions); FED. R. CIV. P. 69 (writs of executions); FED. R. CIV. P. 70 (judgments for specific acts); Feiock v. Feiock, 485 U.S. 624 (1988) (reviewing civil and criminal contempt precedents).

114. Available procedures for certain civil case settlement breaches include criminal contempt, 18 U.S.C. § 401(3) (2000) (disobedience to lawful court order), and compensatory or coercive civil contempt. D. Patrick, Inc. v. Ford Motor Co., 8 F.3d 455, 460 (7th Cir. 1993) (contempt may be used only where breaches involve alleged violations of express and unequivocal commands of court orders). For a review of the forms contempt and suggestions on their use, see Margit Livingston, *Disobedience and Contempt*, 75 WASH. L. REV. 345 (2000).

115. *See, e.g.*, Gompers v. Bucks Stove & Range Co., 221 U.S. 418, 441 (1911).

116. *Id. See also* 18 U.S.C. § 401(3) (criminal contempt includes disobedience to a lawful court order).

side, there may be either coercive civil contempt or compensatory civil contempt.[117] Before there is a contempt proceeding in the settlement breach setting, there usually must be a failure of compliance with an express and unequivocal command within a lawful court order.[118] Thus, contempt may only be available for a settlement breach where the agreement was incorporated into a court order. If the settlement terms were sealed or otherwise outside a court order, but jurisdiction over the settlement was retained, contempt may not be immediately available, though other procedures may be used.[119] Where contempt is available, both civil and criminal proceedings may arise from a single act, though because different procedures apply, they frequently will be presented separately.[120]

A trial court may also proceed on settlement breaches by way of contract dispute resolution. Here, settlement enforcement often follows the routine contract dispute resolution procedures employed to resolve any factual and legal disputes. Yet, the applicable procedures may not always be the same as they would for ordinary contract disputes involving such matters as defective widgets; for example, more "summary" procedures may be appropriate for settlement enforcement.[121]

---

117. Int'l Union v. Bagwell, 512 U.S. 821, 827-29 (1994).

118. *D. Patrick, Inc.*, 8 F.3d at 460. In rare settings, perhaps, breach of an unincorporated settlement agreement may also be misbehavior in the vicinity of the court that obstructs the administration of justice and triggers possible contempt. 18 U.S.C. § 401(1).

119. *See, e.g., D. Patrick, Inc.*, 8 F.3d at 457-58, 462 (suggesting that while contempt procedures were unavailable to enforce an earlier settlement that was not incorporated into a court order, breach of contract procedures could be used because the trial court expressly retained jurisdiction "for the purposes of the enforcement"); Central States S.E. & S.W. Pension Fund v. Richardson Trucking, Inc., 451 F. Supp. 349, 350 (E.D. Wis. 1978) ("Here the orders in both cases are in substance injunctive. However, the orders did not themselves set forth what payments the defendants were required to make, but instead did nothing more than incorporate the terms of the parties' agreements with respect to payment schedules. The orders thus fail to meet the directive of Rule 65(d), and even if they are disobeyed, they may not be made the subject of civil contempt proceedings.").

120. *See, e.g.*, F.J. Hanshaw Enter., Inc. v. Emerald River Dev., Inc., 244 F.3d 1128 (9th Cir. 2001) (civil contempt finding affirmed, but criminal contempt finding reversed because procedural protections were not present).

121. Often, in settlement enforcement settings, "summary" procedures involve resolution without evidentiary hearings. Where necessary procedures entail evidentiary hearings following formal discovery because of disputes over material issues of fact, jury trials may be needed. *Compare* Millner v. Norfolk & W. Ry. Co., 643 F.2d 1005, 1009 (4th Cir. 1981) (when a material dispute arises regarding a settlement agreement, the "trial court must . . . conduct a plenary evidentiary hearing"); Quint v. A.E. Staley Mfg. Co., No. Civ.96-71-B, 1999 WL 33117190, at *1 (D. Me. Dec. 23, 1999) (usually no jury trial right in settlement enforcement proceedings, with FELA claims possibly excepted); Ford v. Cotozems & S. Bank, 928 F.2d 1118, 1121-22 (11th Cir. 1991) (no jury trial right). Summary settlement enforcement and ordinary contract enforcement procedures both differ from contempt procedures that may be employed when settlement orders are

Certain breaches of settlement pacts incorporated into judgments and involving only "the payment of money" seemingly may also be processed through writs of execution under Federal Rule of Civil Procedure 69(a), "unless the court directs otherwise."[122] Here, the procedures follow the practices of "the state in which the district court is held." These writs can involve such remedies as attachment, garnishment, and sequestration.[123] Unlike written federal laws, some written state laws expressly recognize the opportunity for a judgment creditor to choose between different enforcement procedures. For example, the Illinois Marriage and Dissolution of Marriage Act says that terms of a dissolution agreement "set forth in [a] judgment are enforceable by all remedies available for enforcement of a judgment, including contempt, and are enforceable as contract terms."[124]

Choices of applicable procedures are constrained in some settings. Consider, for example, cases where settling parties wish to keep their agreement secret, but nevertheless have the district court retain at least some enforcement jurisdiction. In one recent case, a newspaper sought to intervene in a civil action in order to obtain a copy of such a settlement agreement.[125] The magistrate judge had approved the agreement, but "did not embody his approval in a judicial order that would have made the agreement enforceable by contempt proceedings."[126] The appeals court ruled that such an approval had "no legal significance" to enforcement unless it was "embodied in a judicial order retaining jurisdiction of the case in order to be able to enforce the settlement without a new lawsuit."[127] As to the wish to keep the settlement secret, the appeals court said, "the general rule is that the record of a judicial proceeding is public" and that concealing records disserves the values protected by the First Amendment and bars the public from monitoring judicial performance adequately.[128] The appeals court found there was "a strong presumption," rather than an absolute rule, of

---

disobeyed. *See, e.g., D. Patrick, Inc.*, 8 F.3d at 459 ("because the contempt proceeding is concerned solely with whether or not the respondent's conduct violates a prior court order, the parties cannot reasonably expect to litigate to the same extent that they might in a new and independent civil action"); *F.J. Hanshaw*, 244 F.3d at 1143 n.11 (need finding of bad faith in civil contempt proceeding, perhaps based on clear and convincing evidence).

122. FED. R. CIV. P. 60(a). In "extraordinary circumstances" Fed. R. Civ. P. 70 may be used. *See, e.g.*, Spain v. Mountanos, 690 F.2d 742, 744-45 (9th Cir. 1982) ("under the extraordinary circumstances here where the [money] judgment is against a state which refuses to appropriate funds through the normal process . . . any remedy provided in Rule 69 or Rule 70 to enforce the award" is appropriate).

123. *In re* Merrill Lynch Relocation Mgmt., Inc. v. Merrill Lynch Relocation Mgmt., Inc., 812 F.2d 1116, 1120 (9th Cir. 1987) (Rule 69(a) has been applied "to garnishment, mandamus, arrest, contempt of a party, and appointment of receivers").

124. 750 ILL. COMP. STAT. ANN. 5/502(e) (2002).

125. Jessup v. Luther, 277 F.3d 926, 927 (7th Cir. 2002).

126. *Id.*

127. *Id.* at 929.

128. *Id.* at 927-28.

openness.[129] So upon "a compelling interest in secrecy," the record of an enforceable settlement could be sealed.[130] The court noted most "settlement agreements, like most arbitration awards and discovery materials, are private documents. . . not judicial records," and thus the issue of balancing the interest in promoting settlements by preserving secrecy versus the interest in making public materials upon which judicial decisions are based does not arise.[131] The issue does not arise because there is "no judicial decision" where there is "a stipulation of dismissal . . . without further ado or court action," leaving the settlement with "the identical status as any other private contract."[132] Since the trial judge in the case had participated in "the making of the settlement," the appeals court found the "fact and consequence of his participation are public acts."[133] So, future ancillary enforcement jurisdiction may be unavailable to many parties who wish secrecy for their settlements.

Choices of applicable procedures are also constrained in certain settings where settling parties or their attorneys may later wish to pursue an award of attorney's fees. For example, fees may be awarded to "the prevailing party" in certain civil rights actions.[134] The U.S. Supreme Court has ruled that a determination of "legal merit" is a condition for such an award and that a consent decree may meet this condition if it involves judicial approval and oversight of "court-ordered change in the legal relationship" between the settling parties.[135] One federal court has ruled that such a consent decree arises when a trial court incorporates a settlement into an order, making the contractual obligations enforceable as an order of court, but may not arise when a trial court retains enforcement jurisdiction over a settlement which has not been incorporated.[136]

G. Discretionary Refusals of Settlement Enforcement Requests

Where a federal district court has incorporated terms of a settlement agreement into an order or has retained jurisdiction to enforce a settlement agreement, can it later decline to enforce the settlement even though requested, leaving the matter to other courts? If so, under what circumstances? Or, is such

---

129. *Id.* at 928.

130. *Id.*

131. *Id.* (citation omitted).

132. *Id.*

133. *Id.* at 929.

134. *See, e.g.*, 28 U.S.C. § 1988(b) (1994 & Supp. 1999).

135. Buckhannon Bd. & Care Home v. W. Va. Dep't of Health & Human Res., 532 U.S. 598, 604 (2001) (quoting Tex. State Teachers Ass'n v. Garland Indep. Sch. Dist., 489 U.S. 782, 792 (1989)). The same "prevailing party" standard seemingly operates in other civil rights settings where fee awards are allowed. *See* Race v. Toledo-Davita, 291 F.3d 857 (1st Cir. 2002) (America with Disabilities Act claims); Oil, Chem. & Atomic Workers Int'l Union v. Dep't of Energy, 288 F. 3d 452 (D.C. Cir. 2002) (using standard in fee requests under Freedom of Information Act).

136. *See* Roberson v. Giuliani, 2002 WL 253950, at *6 (S.D.N.Y. Feb. 21, 2002); Smyth v. Rivero, 282 F.3d 268, 285 (4th Cir. 2002).

enforcement exclusively within the subject matter jurisdiction of that district court, so that no other court (federal or state) may enforce? To date there has been little attention to these questions.

We reject the notion of exclusive subject matter jurisdiction in the trial court where the settlement was reached, even where there is an incorporation of the agreement or a retention of jurisdiction. Where enforcement jurisdiction is ancillary, judicial discretion about its exercise should remain available as it does in similar settings, such as when federal district courts are asked to exercise "supplemental" jurisdiction.[137] When a settlement dispute involves "a novel or complex issue of [s]tate law,"[138] federal enforcement jurisdiction often should be declined. Yet, employment of the same standards in enforcement settings that are used in other ancillary jurisdiction settings would be inappropriate. Thus, enforcement should not be declined simply because all claims over which there was original jurisdiction have been dismissed.[139] If the discretion to decline to exercise ancillary enforcement power is used too liberally where the settlement was incorporated into a court order or where jurisdiction was expressly retained, the future settlements will be deterred and certain judicial efficiencies will be undermined. Therefore, there should be very little discretion to refuse enforcement requests where earlier court orders expressly provided for "exclusive" jurisdiction over later disputes.[140]

In addition to at least some of the standards used with statutory supplemental jurisdiction, we posit additional general guidelines on discretionary refusals of settlement enforcement requests. First, refusals should be more difficult where federal law claims were settled because there is a greater likelihood that federal laws will govern legal issues arising during enforcement proceedings. Second,

---

137. 28 U.S.C. § 1367(c) (2000). The extent to which enforcement jurisdiction may be exercised under the supplemental jurisdiction statute remains somewhat unclear. To us, at least some exercise is appropriate under 28 U.S.C. § 1367(a) (allowing supplemental jurisdiction over "claims that are so related to claims in the action within . . . original jurisdiction that they form part of the same case or controversy"). *See* Kokkonen v. Guardian Life Ins. Co. of Am., 511 U.S. 375, 379 (1994) (recognizing that in some instances settlement enforcement claims and claims earlier presented for judicial resolution may have something to do with each other in that they are all "in varying respects and degrees factually interdependent").

138. 28 U.S.C. § 1367(c)(1) (granting court discretion to decline supplemental jurisdiction when "claim raises a novel or complex issue of State law").

139. *But see* 28 U.S.C. § 1367(c)(3) (granting court discretion to decline supplemental jurisdiction when "court has dismissed all claims over which it has original jurisdiction").

140. While parties cannot establish federal district court subject matter jurisdiction by contract, the incorporation of an exclusive venue provision in a court order in a pending civil action signifies a judicial recognition that there will be ancillary jurisdiction in certain events, in addition to providing a judicial promise that, in the absence of exceptional circumstances, it will be exercised. *See, e.g.*, Manges v. McCamish, Martin, Brown & Loeffler, 37 F.3d 221, 224 (5th Cir. 1994). *But see* Housing Group v. United Nat'l Ins. Co., 109 Cal. Rptr. 2d 497 (Ct. App. 2001) (persons involved in settlement talks outside of any civil lawsuit cannot agree to place settlement before a trial court in order to secure possible court enforcement because there is no justiciable controversy).

refusals should be more difficult where the same district judge will preside over the settlement enforcement proceedings as presided over the settlement talks because desired efficiencies are more likely to occur.[141] Third, refusals should be easier when federal governmental interests are diminished due to settlement agreements which expressly require that state laws govern any future disputes. Fourth, refusals should be more difficult where enforcement proceedings will involve settlement breaches that violate court orders because they more readily implicate the power of the courts to "protect" their proceedings and to "vindicate" their authority.[142] Fifth, refusals should be easier where enforcement proceedings will not involve extensive inquiries into court records, such as hearing transcripts and filed papers. Sixth, refusals should be more difficult where earlier and related settlement enforcement proceedings have already occurred in the federal district court.

III. IMPROVING SETTLEMENT ENFORCEMENT IN THE FEDERAL DISTRICT COURTS

Many of the difficulties with federal settlement enforcement proceedings can be reduced by new written federal laws. We posit that such new laws are needed both from the U.S. Supreme Court, as the federal civil procedure rulemaker, and from the Congress. As rulemaker, the Court should consider both amendments to existing civil procedure rules and entirely new rules. We urge Congress at this time to focus only on changes to the supplemental jurisdiction statute.

Difficulties regarding the incorporation of settlement terms into court orders and the retention of jurisdiction for later enforcement could be reduced through amendments to Federal Rule of Civil Procedure 58. The rule already speaks to judgments upon jury verdicts or other decisions by juries, as well as to judgments upon decisions by courts without juries.[143] An amended rule could be accompanied by new forms, which would reduce confusion, as they would be "sufficient" if used.[144] An amended rule could be modeled on some existing state civil procedure laws. For example, a Texas statute says:

> (a) If the parties reach a settlement and execute a written agreement disposing of the dispute, the agreement is enforceable in the same manner as any other written contract.
> (b) The court in its discretion may incorporate the terms of the agreement in the court's final decree disposing of the case.
> (c) A settlement agreement does not affect an outstanding court order unless the terms of the agreement are incorporated into a subsequent decree.[145]

---

141. *Kokkonen*, 511 U.S. at 380 ("efficient" to adjudicate settlement breach with claim prompting the settlement where facts underlying both have much "to do with each other").

142. *Id.* at 380-81.

143. FED. R. CIV. P. 58.

144. FED. R. CIV. P. 84 (forms in Appendix of Forms are sufficient).

145. TEX. CIV. PRAC. & REM. §154.071.

And, a California Code of Civil Procedure says:

> If parties to pending litigation stipulate, in a writing signed by the parties outside the presence of the court or orally before the court, for settlement of the case, or part thereof, the court, upon motion, may enter judgment pursuant to the terms of the settlement. If requested by the parties, the court may retain jurisdiction over the parties to enforce the settlement until performance in full of the terms of the settlement.[146]

Difficulties regarding discretionary refusals of future or present settlement enforcement requests could be reduced through amendments to the supplemental jurisdiction statute.[147] That statute is applied today, for the most part, to the initial adjudicatory authority over civil claims pleaded or otherwise presented before or during so-called trials on the merits, typically encompassing "factually interdependent" claims under *Kokkonen*.[148]

Further difficulties with settlement enforcement procedures can be diminished with amendments to Federal Civil Procedure Rules 65 and 69. Amendments to Federal Rule of Civil Procedure 65(d) could address enforcement issues arising from settlements involving equitable remedies. Amendments to Federal Rules of Civil Procedure 69(a) could address enforcement issues arising from settlements involving monetary payments. Should codification of civil contempt procedures be found necessary, a new federal civil procedure rule seems the best vehicle to do so[149] using several local court rules and written state laws as models.[150]

CONCLUSION

Settlements of federal civil actions may, but need not, be subject to later judicial enforcement. As recognized by the U.S. Supreme Court in *Kokkonen v. Guardian Life Insurance Co.*, one significant limitation on enforcement proceedings is subject matter jurisdiction because federal district courts are "courts of limited jurisdiction." Under *Kokkonen*, enforcement jurisdiction may be "independent," but usually is "ancillary" because state law claims typically are

---

146. CAL. CIV. PRO. CODE §664.6 (1987 & Supp. 2002). Prior to its enactment, "California appellate decisions were in conflict as to the appropriate procedure for enforcement of an agreement to settle pending litigation." Assemi v. Assemi, 872 P.2d 1190, 1194-95 (Cal. 1994). *But see* LA. CIV. CODE ANN. art. 3071 (1994) (settlement recited in open court "confers" upon each party "the right of judicially enforcing its performance").

147. 28 U.S.C. § 1367 (2000).

148. A review and critique of the supplemental jurisdiction statute appears in Jeffrey A. Parness & Daniel J. Sennott, *Expanded Recognition in Written Laws of Ancillary Federal Court Powers: Supplementing the Supplemental Jurisdiction Statute*, ___ U. PITT. L. REV. ___ (2002).

149. Acts constituting criminal contempt are already expressly addressed in 18 U.S.C. § 401 (2000). These statutory standards have traditionally been used to help define acts constituting civil contempt.

150. *See, e.g.*, ILL. CIR. CT. R. FOR FIFTEENTH CIR. 11.1 (2000); CONN. SUP. R. § 1-14 (1999).

involved where there is no diversity of citizenship. Ancillary enforcement powers may be exercised by district courts either where claims were initially presented for adjudication and disputes arising from later settlements are "factually interdependent," or where recognition of enforcement authority enables courts "to function successfully," such as where courts need to insure that their orders are not "flouted or imperiled." Typically, enforcement authority is exercised so that the courts function successfully.

Difficulties have surfaced regarding this ancillary settlement enforcement jurisdiction. They concern how to incorporate settlement terms into court orders and how otherwise to retain jurisdiction, whether settlement disputes may prompt the reopening of judgments, and what contract laws and what procedures should apply when federal case settlements are enforced. There are additional troubles which have yet to surface significantly, including whether there is judicial discretion to refuse requests that future enforcement jurisdiction be retained and whether certain settlement disputes can prompt discretionary refusals of available enforcement jurisdiction.

We believe new written federal laws are needed now to address many of these difficulties. Relevant lawmakers include both the U.S. Supreme Court, as promulgator of the federal rules of civil procedure, and the Congress. We suggest amendments to the Federal Rules of Civil Procedure on judgment entry, on judgments involving money and on permanent injunctions, as well as changes to the supplemental jurisdiction statute.

AN INTERPRETATION AND (PARTIAL) DEFENSE OF LEGAL FORMALISM

PAUL N. COX*

INTRODUCTION

The origin of this lecture lies in an observation. Specifically, I was struck by a substantial similarity in the views of Grant Gilmore and of Friedrich Hayek. What is striking in this observation is that Gilmore was a kind of legal realist. As a realist his skepticism about law was expressed as an attack upon legal formalism.[1] Hayek, by contrast, is at least generally characterized as a legal formalist.[2] And what I view as Hayek's very similar skepticism about law was expressed as advocacy of legal formalism.

What is the nature of the skepticism that I, at least, view as common to both of these eminent legal thinkers? At bottom, it is, both distrust of and distaste for centralized, all encompassing legal direction. Gilmore put it this way:

> As lawyers we will do well to be on our guard against any suggestion that, through law, our society can be reformed, purified, or saved. The function of law, in a society like our own, is altogether more modest and less apocalyptic. It is to provide a mechanism for the settlement of disputes in the light of broadly conceived principles on whose soundness, it must be assumed, there is a general consensus among us.[3]

Repeatedly in his work, Hayek makes what I believe is a substantially similar point: "constructivist rationalism," the belief that, by means of a "scientific" law, society may be purposefully reconstructed, and human activity directed to serve collectively determined goals, is a tragically false, dangerous and destructive myth.[4] Gilmore identifies formalism with that myth. Hayek offers formalism as

---

* Centennial Professor of Law, Indiana University School of Law—Indianapolis. This inaugural lecture was delivered on March 7, 2002, at the Indiana University School of Law—Indianapolis.

1. *See* GRANT GILMORE, THE AGES OF AMERICAN LAW 41-67 (1977).

2. *E.g.*, MORTON J. HORWITZ, THE TRANSFORMATION OF AMERICAN LAW, 1870-1960, at 228-30 (1992); RICHARD A. POSNER, THE PROBLEMS OF JURISPRUDENCE 57, 60 (1990).

3. GILMORE, *supra* note 1, at 109.

4. *See* 1 F. A. HAYEK, LAW, LEGISLATION AND LIBERTY: RULES AND ORDER 8-34 (1973) [hereinafter HAYEK, LAW, LEGISLATION AND LIBERTY]; 1 F. A. HAYEK, *Traditional Morals Fail to Meet Rational Requirements*, *in* THE FATAL CONCEIT: THE ERRORS OF SOCIALISM 66-88 (W.W. Bartley, III ed., 1988); FRIEDRICH A. HAYEK, THE CONSTITUTION OF LIBERTY 234-49 (1960) [hereinafter HAYEK, CONSTITUTION OF LIBERTY]. Whether Hayek is properly characterized as a formalist is debatable. His views on law changed from the time of the more clearly formalist *Constitution of Liberty*, HAYEK, CONSTITUTION OF LIBERTY, *supra*, to the time of *Law, Legislation and Liberty*, HAYEK, LAW LEGISLATION AND LIBERTY, *supra*. The change is attributable to the influence on Hayek of Leoni. *See* BRUNO LEONI, FREEDOM AND THE LAW (3d ed., Liberty Fund, Inc. 1991).

an alternative to and defense against the myth.

Who was right? For me, the question is particularly interesting because I was brought up in the law to believe that formalism is a sin. This is not an experience unique only to me. It is, I venture to guess, an article of faith among most legal academics that formalism is a sin—which is not to say that formalism is absent from contemporary law, or even from contemporary academic commentary. Indeed, judging from that commentary, there is far too much formalism going on. For formalism, as a sin, is the label the commentators often attach to the targets of their critique.[5] A difficulty with this attaching of that label is that the precise content of the sin supposed to have been committed is often unclear.

What is legal formalism?

As formalism is most often defined by its critics,[6] and as the critics often have arguably distinct targets in mind, the question is perhaps better framed as "what are legal formalisms?" At least this is so unless there is some underlying foundational belief at the bottom of the variety of formalisms, one that implies or necessitates each.

In surveying the various legal formalisms, I will rely in part upon positions taken or said to have been taken by the "classical formalists"—legal academics writing at the end of the Nineteenth Century and beginning of the Twentieth Century, who were principally associated with the Harvard Law School, and with the then dean of that school, Christopher Columbus Langdell.[7] However, I am not engaged in an exercise of legal history, and I am not, therefore, seeking to recapture the particulars of the thought of these academics. Rather, I am both outlining contemporary beliefs about what formalism is or was, whether or not these contemporary beliefs accurately portray the long lost era of classical formalism, and constructing an interpretation of the formalist impulse, one only partially related to the specifics of classical formalism.

Similarly, I will refer to formalism's critics as legal realists, post-realists or

---

5. *See, e.g.*, RICHARD A. POSNER, OVERCOMING LAW 19-20, 75-76 (1995) [hereinafter POSNER, OVERCOMING LAW]; POSNER, *supra* note 2, at 274-75; CASS R. SUNSTEIN, LEGAL REASONING AND POLITICAL CONFLICT 24-26 (1996).

6. *See* Frederick Schauer, *Formalism*, 97 YALE L.J. 509 (1988) (making this point in context of a defense of formalism). However, there has been a recent renewal of interest in formalism, and there are contemporary defenders of various varieties of formalism. *See, e.g.*, ERNEST J. WEINRIB, THE IDEA OF PRIVATE LAW (1995); Michael Corrado, *The Place of Formalism in Legal Theory*, 70 N.C. L. REV. 1545 (1992); Schauer, *supra*; Alan Schwartz, *Incomplete Contracts*, 2 NEW PALGROVE DICTIONARY OF ECONOMICS AND LAW 277 (1997); Ernest J. Weinrib, *Legal Formalism: On The Immanent Rationality of Law*, 97 YALE L. J. 949 (1988); James G. Wilson, *The Morality of Formalism*, 33 UCLA L. REV. 431 (1985); *see also* Symposium, *Formalism Revisited*, 66 U. CHI. L. REV. 527 (1999) (exploring contemporary relevance of varieties of formalism).

7. For contemporary depictions of the classical formalists, see, e.g., NEIL DUXBURY, PATTERNS OF AMERICAN JURISPRUDENCE 9-64 (1995); GARY MINDA, POSTMODERN LEGAL MOVEMENTS 13-33 (1995); ROBERT SAMUEL SUMMERS, INSTRUMENTALISM AND AMERICAN LEGAL THEORY 136-59 (1982).

pragmatic instrumentalists.[8] I am aware that legal realism was less a coherent school of thought than a set of somewhat diverse impulses, but I am not presently interested in the details of legal realism, the differences between particular legal realists or the differences between legal realism and the post-realist schools that incorporate realist insights. Realism, post-realism, and pragmatic instrumentalism are largely employed here merely as labels for anti-formalist arguments and positions. Nevertheless, it will become apparent that I offer an interpretation of the "realist" impulse, just as I do of the formalist impulse.

My objective is a reconstruction of formalism on grounds of skepticism about legal competence. This will strike many as a peculiar, even perverse thesis. A common theme in anti-formalist thought is precisely that formalism entails an exaggerated, and erroneous, belief in legal competence, it is a belief that the formalist legal method is adequate to the task of properly resolving problems confronted in law.[9] I do not deny that formalist rhetoric often appears imperious, but I offer an interpretation of formalism that depicts it as devoted to a constrained ambition for law. In the course of my survey of legal formalisms, I will also identify what I take to be the principal objections to the formalism in question, and I will suggest at least partial rebuttals. I proceed initially in three parts, addressing, in turn, formalism as autonomous conceptualism, formalism as rules, and formalism as empty spaces. I then seek to address the merits of formalism and its chiefly consequentialist competitors.

I. FORMALISM AS AUTONOMOUS CONCEPTUALISM

What is "autonomous conceptualism"? By "autonomous" I mean that at least classical formalists believed that answers to legal questions could and should be based upon distinctly legal materials, without reference to sources external to

---

8. I therefore employ the term "legal realist" in a very broad sense in this essay to include not merely the legal realists of the 1930s, but proto-realists, such as the early Roscoe Pound, and post-realists. Post-realists include all who would agree with the claim that "we are all realists now" in the sense that they are committed to what Professor Summers calls "pragmatic instrumentalism." *See* Robert S. Summers, *Pragmatic Instrumentalism in Twentieth Century American Legal Thought—A Synthesis and Critique of Our Dominant General Theory About Law and Its Use*, 66 CORNELL L. REV. 861 (1981). I *exclude* from "legal realism" as I employ the phrase, that branch of legal realism devoted to extreme skepticism or nihilism. So "realism" in my usage refers to the pragmatic, social science branch of the phenomenon.

9. This is obviously apparent in Gilmore, but it was also a common theme in legal realist literature and is a theme in Judge Posner's critique of contemporary legal practice. *See, e.g.*, POSNER, OVERCOMING LAW, *supra* note 5; Felix S. Cohen, *Transcendental Nonsense and the Functional Approach*, 35 COLUM. L. REV. 809 (1935); Duncan Kennedy, *The Structure of Blackstone's Commentaries*, 28 BUFF. L. REV. 205 (1979); David Lyons, *Legal Formalism and Instrumentalism—A Pathological Study*, 66 CORNELL L. REV. 949 (1981); Gary Peller, *The Metaphysics of American Law*, 73 CAL. L. REV. 1151 (1985). On the other hand, some recent "formalist" proposals are predicated on the idea that formalism may be the best that can be done given the incapacities of legal actors. *E.g.*, Eric Posner, *A Theory of Contract Law Under Conditions of Radical Judicial Error*, 94 N.W. U. L. REV. 749 (2000).

law, most obviously without reference to the social sciences.[10] By "conceptualism," I mean that at least classical formalists believed three things.[11] First, legal concepts, such as the concept of consideration in contract or the concept of ownership in property, could be identified through *induction*, though that is a review of the evidence of case law. Second, they believed that more particular rules could then be derived "logically" from the concepts induced from the caselaw. Third, they believed that the result would be a self-contained, internally consistent, systemized and rationalized law, rather like geometry, and, therefore, that correct legal answers could be given to any question by reference to the logic of this system.

This, at least, is the standard account, the account attacked by Holmes[12] and later by legal realists.[13] What, then, is wrong with autonomous conceptualism? I will not review all of the criticisms, but I will attempt a summary of the main lines of attack. First, the concepts employed by the classical formalists were far too general. The radical version of this criticism was a nominalist belief that concepts do not have real world referents, or that real world referents are insufficiently identical to be captured by any concept.[14] A more moderate version of the criticism is that only narrow concepts drawn at lower levels of abstraction can be serviceable for formalist law.[15] Thus, for example, abstract concepts like "ownership" or "property right" or "liberty" cannot yield particular uncontroversial legal conclusions because various possible conclusions may follow from them. In Hohfeldian terms, abstract concepts such as property must be disaggregated before they become descriptive of the actual variety of possible legal relationships.[16] An implication of this view is that judges are not in fact

---

10. *See* Thomas C. Grey, *Langdell's Orthodoxy*, 45 U. PITT. L. REV. 1, 16-20 (1983). "Formalism" would therefore seem to entail one of the central claims of legal positivism: that law is distinct from morality. At least this would seem to be the case if morality means "everything else." Frederick Schauer & Virginia Wise, *Legal Positivism As Legal Information*, 82 CORNELL L. REV. 1080 (1997).

11. *See* Stanley Fish, *The Law Wishes to Have a Formal Existence*, *in* THE FATE OF LAW 159 (Austin Sarat & Thomas Kerns eds., 1991) (offering a somewhat parallel account of formalism, but attributing it to contemporary legal practice); Grey, *supra* note 10.

12. *See* OLIVER WENDELL HOLMES, *The Path of the Law*, *in* COLLECTED LEGAL PAPERS 167 (Peter Smith ed., 1952) (1920). Gilmore nevertheless attacked Holmes as a formalist. *See* GILMORE, *supra* note 1, at 48-56. In terms of this essay, Holmes is best viewed as a proto-realist in his (moderate) attack on formalism as autonomous conceptualism and as a formalist in his preference both for rules and for empty spaces. *See generally* DUXBURY, *supra* note 7, at 37-47; Grey, *supra* note 10, at 44.

13. *E.g.*, JEROME FRANK, LAW AND THE MODERN MIND (Peter Smith ed., Anchor Books 1970) (1930); Cohen, *supra* note 9; John Dewey, *Logical Method and Law*, 10 CORNELL L. Q. 17 (1924); Roscoe Pound, *Mechanical Jurisprudence*, 8 COLUM. L. REV. 605 (1908).

14. *See* AMERICAN LEGAL REALISM 166 (William W. Fisher, III et al. eds., 1993).

15. CASS SUNSTEIN, LEGAL REASONING AND POLITICAL CONFLICT 24-26 (1996).

16. *Id.*; *see* Wesley Hohfeld, *Some Fundamental Legal Conceptions as Applied in Judicial Reasoning*, 23 YALE L.J. 16 (1913).

bound by concepts, as these may be manipulated.[17] If particular rules or rights are not in fact compelled by the high level abstractions relied upon by formalists, judges are not in fact engaged in finding the law and following it. Rather, they are engaged in willing the results they reach in the particular cases they decide.

Second, and perhaps more importantly, formalism's geometrical aspirations are normatively suspect. What is needed instead, said Holmes, the realists, the pragmatists, and most recently Judge Posner, is a concrete focus upon considerations of social advantage and disadvantage.[18] Legal decision should not proceed then from fidelity to the heaven of legal concepts, but rather from consideration of the consequences of alternative decisions. Law, in this anti-formalist depiction, is an instrument of social policy to be used for socially desirable ends. An implication of this normative critique of formalism is denial of law's autonomy: if law is an instrument to be purposively applied, it requires the tools and information supplied by "science" of one sort or another.

These, I think, summarize the main lines of attack, but there is a third line, distinct from and arguably antagonistic to the second, a line most obviously associated with Karl Llewellyn: abstract formalist concepts should be replaced with context dependent sensitivity to social practice.[19] Law should be specific to situation types or categories and should incorporate the norms of real people in the real world. It should be noticed that this reference to social practice as a source of law has much in common with Hayek's Humean theory of spontaneous order and with, at least at some points in Hayek's intellectual journey, his recommendations for law.[20] It may also be a point of partial commonality between Hayek and Gilmore. However, there is a tension between the second and this third critique of autonomous conceptualism in *at least* one respect: the preferred source of law in the second is science; the preferred source in the third is practice.

What might be said of formalism given these critiques? I cannot defend formalism in its pristine, classical sense for two reasons. First, it is simply not an accurate depiction of law as it now is, even if, which is doubtful, it once was such a depiction. I would be guilty of malpractice if I described our law in classically formalistic terms and if I taught it in these terms. Second, I think the critique of generalized abstraction partially correct: legal particulars cannot be

---

17. *Cf.* John Harrison, *The Power of Congress to Limit the Jurisdiction of the Federal Courts and the Text of Article III*, 64 U. CHI. L. REV. 203, 253 (1997) (explaining that principles compete with each other and any given principle can be implemented in a variety of ways).

18. *See generally supra* notes 9, 13. For one of Judge Posner's recent statements, see POSNER, OVERCOMING LAW, *supra* note 5, at 399.

19. *See* KARL N. LLEWELLYN, THE COMMON LAW TRADITION, DECIDING APPEALS 127 (1960); WILLIAM TWINING, KARL LLEWELLYN AND THE REALIST MOVEMENT chs. 11-12 (1973).

20. *See* HAYEK, LAW, LEGISLATION AND LIBERTY, *supra* note 4, at 35-54, 74-91, 100-01; Symposium, *Decentralized Law for a Complex Economy*, 23 SW U. L. REV. 443 (1994); Symposium, *Public and Private Ordering and the Production of Legitimate and Illegitimate Rules*, 82 CORNELL L. REV. 1123 (1997). Indeed, Hayek in his later work attacks Langdellian versions of autonomous conceptualism. *See* HAYEK, LAW, LEGISLATION AND LIBERTY, *supra* note 4, at 105-06.

uncontroversially derived from abstract concepts, and the law is unlikely ever to achieve a state of internal consistency.

Nevertheless, I wish to offer a partial defense of autonomous conceptualism. My initial point is that a substantial degree of conceptualism is inescapable in law, and a substantial degree of conceptualistic argument is evident in law. Conceptualism is inescapable because one does not, contrary to the view of some realists, approach facts without reference to concepts and expect to do anything intelligible.[21] Concepts are essential to thought about and evaluation of facts; recognition of this fact should lead to a preference for making one's concepts explicit. Moreover, conceptualism is normatively essential. The nominalist's rejection of conceptual ordering generates radical case specific decision: if no two cases are sufficiently alike to justify a concept or rule encompassing them, there can be no such concept or rule. This is a formula for rule by arbitrary prejudice, not law.

That there is a substantial degree of conceptualistic argument in law is evident not only in any casual reading of appellate opinions, but also in contemporary legal theory. Dworkin, in substituting "equality" for "liberty," "fit" for "deduction" and "moral philosophy" for "existing case law" may be demonstrating a more sophisticated technique than Langdell, but his remains a species of conceptualism.[22] Neoclassical economic analysis of law is obviously a formalist enterprise in its technique: through deduction from the rationality and scarcity postulates it generates hypotheses, which hypotheses are then formulated as legal rules. True, the object of this enterprise is consequentialist: it is not, or is not supposed to be, undertaken as an act of fidelity to rationality and scarcity, but as an instrument for identifying social advantage understood as efficiency.[23] On the other hand, to the extent that its hypotheses are unverified or unverifiable, it operates as formalism in precisely the sense that it exhibits a strict fidelity to rationality and scarcity.[24] What, of course, distinguishes these examples from classical autonomous conceptualism is that neither adopt purely legal materials as bases for their conceptualism.

A second point I wish to make in defense of autonomous conceptualism is that the debate between formalists and realists entails, at bottom, a striking difference in perspective over the role of law and the competence of law givers and appliers. Consider in particular the formalist claim that legal particulars are derived from and bound by preexisting concepts and the realist claim that law is an instrument for achieving social purposes.

---

21. *See* L.L. Fuller, *American Legal Realism*, 82 U. PA. L. REV. 429, 443-47 (1934).

22. *E.g.*, RONALD DWORKIN, A MATTER OF PRINCIPLE (1985); RONALD DWORKIN, LAW'S EMPIRE (1986); Ronald Dworkin, *In Praise of Theory*, 29 ARIZ. ST. L.J. 353 (1997). *See* RICHARD POSNER, THE PROBLEMATICS OF MORAL AND LEGAL THEORY 92-120 (1999) (criticizing Dworkin's moral conceptualism).

23. POSNER, OVERCOMING LAW, *supra* note 5, at 17-19.

24. A common complaint leveled at economic analysis is that it is insufficiently supported by empirical evidence. I would argue that, even where supported, the support is often ambivalent, subject to challenge or otherwise inconclusive. *See infra* notes 130-37 and accompanying text.

I will approach these claims through an example. I think it fair to say that a limited, bargain view of contract, a view requiring exchange of consideration to achieve legal enforceability, was a formalist notion.[25] The effect of the notion, consistently applied, was to deny enforcement to many promises and, in particular, to largely deny legal protection to reliance interests. These consequences followed from a derivation of particular rules from the concept of bargain.[26] By contrast, realist and post-realist contract law either rejects or extends the bargain principle so as both to enforce more promises and to provide a measure of protection to reliance interests.[27] It does so, in realist fashion, by contending that the purposes of the bargain principle are better served by expanding or ignoring it, or by contending that the harms generated by inducing reliance are worthy of legal protection[28].

At one level of analysis this example illustrates the distinction between a rigid deduction of legal result from abstract concept in formalist law and the treatment of law as a purposive instrument for achieving ends (for example, the end of encouraging exchange) in realist and post-realist law. Consider, however, a further level: the formalist's adherence to the bargain principle served the end of freedom *from* legal enforcement of promises, that is, freedom *from* contract. The realist's position serves the end of freedom *to* contract in the sense that it facilitates the practice of effective promise making. The costs of the realist's position, however, are that it requires a substantially greater role for the governmental functionary known as the judge and relies upon a questionable assumption about the competence of that judge, for enforcement of promises beyond the original limits of the bargain principle requires either a difficult empirical inquiry into the seriousness of an often ambiguous promise or the imposition of a tort-like obligation on the basis of the court's perception of proper behavior.[29] Gilmore, recognizing this, declared "The Death of Contract."[30] My difficulty, not Gilmore's, with the expansion of enforceable promise is that it assumes a greater competence in the judge, or judge and jury, than I think warranted.[31] To the extent that what is in issue is what was meant or

---

25. W. DAVID SLAWSON, BINDING PROMISES, THE LATE 20TH CENTURY REFORMATION OF CONTRACT LAW, ch. 1 (1996); Melvin Aron Eisenberg, *The Principles of Consideration*, 67 CORNELL L. REV. 640 (1982).

26. Eisenberg, *supra* note 25, at 641-56.

27. *E.g.*, Lon Fuller & William Perdue, Jr., *The Reliance Interest in Contract Damages: 2*, 46 YALE L.J. 373, 418-20 (1937).

28. Eisenberg, *supra* note 25, at 641-56. *See* Richard Posner, *Gratuitous Promises in Law and Economics*, 65 J. LEGAL STUD. 411 (1977).

29. Jay Feinman, *Promisory Estoppel and Judicial Method*, 97 HARV. L. REV. 678, 712-16 (1984).

30. GRANT GILMORE, THE DEATH OF CONTRACT (1974).

31. *Cf. id.* at 52-54 (explaining contradiction between bargain theory of contract and absolute liability potentially as effort to limit litigation); Richard Craswell, *Offer, Acceptance and Efficient Reliance*, 48 STAN. L. REV. 481, 544-53 (1996) (recognizing problems of unpredictable results from case by case assessments of efficient reliance, but ultimately rejecting bright line rule alternative).

reasonably understood, the highly stylized, long after the fact and frankly largely bizarre performance art we call the trial is an implausible procedure for determining that question. To the extent that the issue is one of the relative costs and benefits, the notion that these can be quantified and compared "objectively" after the fact strikes me as absurd.[32]

My point is this: formalist conceptualism served the end of limiting the scope of law in the sense that it limited occasions on which legal functionaries would assess conduct and therefore occasions on which persons would be called upon to justify their actions before such functionaries. The realist and post-realist ambition, by contrast, is the expansion of these occasions. This should not be surprising; it is inherent in the anti-formalist's treatment of law as an instrument for achieving social purposes. That treatment postulates a collective purpose or collectively determined end state as an objective, an organic beneficiary of this end-state and someone, presumably the legal functionary, as the formulator and implementor of the objective.[33] The obvious questions, ones I will return to at the end of this essay, are whether there is an adequate means of establishing any such objective and whether any such legal functionary can claim sufficient competence in implementation.

Before leaving the matter of autonomous conceptualism, I want to return to the third objection to it, the notion that social practice, rather than abstract formalist concepts should govern law. I wish to make two points about this claim: First, it is not apparent, or, at least, as apparent as realists in Llewellyn's camp believed it to be, that formalist concepts are divorced from social practice. Second, direct resort to social practice is itself fraught with difficulties.

I begin by asking where formalist concepts come from. In Langdellian classical formalism they came from existing case law: the formalist induced them from the practices of the courts.[34] Where, however, did the practices of the courts come from? Langellians apparently didn't ask themselves this question, but let me ask it. One possibility is that it came from some well worked out ideology or moral theory, so the courts were following the precepts of a

---

32. The chief problem with such an objective comparison is the subjectivity of cost. JAMES BUCHANAN, COST AND CHOICE: AN INQUIRY IN ECONOMIC THEORY (1969); F. A. HAYEK, *Economics and Knowledge*, *in* F. A. HAYEK, INDIVIDUALISM AND ECONOMIC ORDER 33 (1948). For discussions of the implications of subjectivity, see, e.g., Gary Lawson, *Efficiency and Individualism*, 42 DUKE L.J. 53 (1992); Gregory Keating, *Reasonableness and Rationality in Negligence Theory*, 48 STAN. L. REV. 311, 337-41, 367-73 (1996). For further discussion of this point, see *infra* note 132 and accompanying text.

33. The contrasts between classical, perhaps formalist law and the post-New Deal administrative state are well depicted in the following: Norman Barry, *The Classical Theory of Law*, 73 CORNELL L. REV. 283 (1988); Donald Gjerdingen, *The Politics of the Coase Theorem and Its Relationship to Modern Legal Thought*, 35 BUFF. L. REV. 871 (1986); and Jerry Mashaw, *"Rights" in the Federal Administrative State*, 92 YALE L.J. 1129 (1983); *cf.* BRUCE ACKERMAN, PRIVATE PROPERTY AND THE CONSTITUTION (1977) (discussing ordinary observer versus scientific policymaker).

34. Grey, *supra* note 10, at 24-27.

Nineteenth Century Ronald Dworkin. Herbert Spencer is, I suppose, a candidate.[35]

That is a possibility, but let me postulate a second one: "intuition." By intuition I mean a set of often tacit commitments, a moral sense, grounded in the "shared morality of a particular society."[36] I think this a possibility for the obvious reason that common law judges of the formalist era were the products of the American society in which they worked. It would be surprising in the extreme if they came up with conclusions, including conclusions consistent with the principles formalists then induced from these conclusions, alien to the conventional understandings and traditions of that society.

This does not mean that formalist adjudications enjoyed or could enjoy universal support from the members of American society, even in the formalist era. It means only that the concepts had some substantial relation to practice. For example, the concept of bargain could be inferred from the actual practice of exchange, and, as a further example, the distinction between act and omission, surely a part of common morality,[37] would, in contrast to strictly consequentialist recommendations, be reflected in law. Nor does it mean that formalist concepts or the rules derived from them tracked in detail actual norms or practices. They would not do so for the reason that norms are inevitably and necessarily distorted if incorporated in law. This is because the addition of legal enforcement to non-legal means of norm enforcement will alter the cost/benefit calculation of the actors subject to the norms, because the mere fact of legal enforcement alters the meaning of norms and because considerations of judicial capacity and administrative cost will often dictate alterations of norms.[38]

---

35. Professor Grey rejects this possibility. *Id.* at 33-35. *Compare* HERBERT HOVENKAMP, ENTERPRISE AND AMERICAN LAW 1836-37, 174-75 (1981) (rejecting connection between classical formalists and *Lochner*), *with* MORTON HORWITZ, THE TRANSFORMATION OF AMERICAN LAW 1870-1960 (1992) (generally making this connection). *See* NEIL DUXBURY, PATTERNS OF AMERICAN JURISPRUDENCE 25-32 (1995) (treating Spencer as source of judicial formalism).

36. Professor Grey raises but rejects this possibility. Grey, *supra* note 10, at 23-24. Nevertheless, it seems to me both that the classical formalist's effort to systemize the common law would necessarily incorporate social custom given an assumption that common law rests upon custom or convention. *E.g.*, MELVIN ARON EISENBERG, THE NATURE OF THE COMMON LAW, Ch. 4 (1988); A.W.B. SIMPSON, *The Common Law and Legal Theory*, *in* OXFORD ESSAYS IN JURISPRUDENCE 77-79 (A.W.B. Simpson ed. 1973). *Cf.* Grey, *supra* note 10, at 30 (evolutionary views of classical formalists rested in part on historical school and therefore upon evolving custom). Moreover, formalism more generally understood entails claims to roots in the historical experience of a people or nation. M. H. Hoeflich, *Law and Geometry: Legal Science from Leibniz to Langdell*, 30 AM. J. LEGAL HIST. 95 (1986). To the extent that the Hayek of RULES AND ORDER, *supra* note 4, can be said to have adopted the common law preferences of Leoni, perhaps his "formalism" entailed an exercise of "finding law" in "existing social-institutional arrangements." *See* James Buchanan, *Good Economics, Bad Law*, 60 VA. L. REV. 483, 488-89 (1974).

37. LEO KATZ, ILL GOTTEN GAINS: EVASION, BLACKMAIL, FRAUD AND KINDRED PUZZLES OF THE LAW (1996).

38. *E.g.*, Randy E. Barnett, *The Sounds of Silence: Default Rules and Contractual Consent*,

Notice that these points raise a question about the desirability of Llewellyn's program, the program of a more direct and concrete incorporation of norms in law, than is suggested by my intuitionist account of formalist principle. A substantial reason for such incorporation is that promises greater degrees of predictability—surely a formalist value.[39] But, if incorporation is inevitably also distortion, the incorporation strategy is problematic. Indeed, it may be that a legal takeover of the norms and understandings of social practice is not what rational persons would prefer. Professor Bernstein has produced at least evidence that they prefer that a rigid, formal and even inequitable law stand outside these understandings as a last resort, leaving adjustment, interpretation and enforcement to non-legal mechanisms of interaction.[40] This is in part because legal enforcement is more costly than its alternative, in part because legal enforcement undermines the alternatives and in part because even the best judges are not competent discoverers of the complexities and often tacit dimensions of social practice. Alternatively, it is because norms are often local affairs and therefore differ between local communities.[41] Inter-local interactions therefore require resolutions that supplant competing local norms.

Llewellyn's critique of formalism may be understood as the claim that formalism divorces law from life, rendering law an alien, unpredictable, and, by reference to the baseline of social practice, arbitrary force.[42] Perhaps, but there is another way of looking at this matter. The question is what version of law, the formalist version or the anti-formalist, instrumental version, poses the greatest threat to life outside it? Llewellyn's attempt to protect life from law through incorporation of life's norms into law can be seen as in fact a threat to life if the

---

78 VA. L. REV. 821, 908 n.231 (1992); Charles Goetz & Robert Scott, *The Limits of Expanded Choice: An Analysis of the Interactions Between Express and Implied Contract Terms*, 73 CAL. L. REV. 261, 275-76 (1985); Richard Pildes, *The Destruction of Social Capital Through Law*, 144 U. PA. L. REV. 2055 (1996); Alan Schwartz, *Relational Contracts in the Courts: An Analysis of Incomplete Agreements and Judicial Strategies*, 21 J. LEGAL STUD. 271 (1992); Alan Schwartz, *The Default Rule Paradigm and the Limits of Contract Law*, 3 S. CAL. INTERDISC. L.J. 389, 404-06 (1993).

39. Fuller, *supra* note 21, at 431-38 (describing Llewellyn's views).

40. *See, e.g.*, Lisa Bernstein, *Merchant Law in a Merchant Court: Rethinking the Code's Search for Immanent Business Norms*, 144 U. PA. L. REV. 1765 (1996); Lisa Bernstein, *The Questionable Empirical Basis of Article 2's Incorporation Strategy: A Preliminary Study*, 66 U. CHI. L. REV. 710 (1999) [hereinafter Bernstein, *Questionable Empirical Basis*]; David Charny, *Non-Legal Sanctions in Commercial Relationships*, 104 HARV. L. REV. 373 (1990); Edward Rock & Michael Wachter, *The Enforceability of Norms and the Employment Relationship*, 144 U. PA. L. REV. 1913 (1996); Robert E. Scott, *The Case for Formalism in Relational Contract*, 94 NW. U. L. REV. 847 (2000).

41. Bernstein, *Questionable Empirical Basis*, *supra* note 40; David Charny, *The New Formalism in Contract*, 66 U. CHI. L. REV. 842 (1999); Richard A. Epstein, *Confusion About Custom: Disentangling Informal Customs from Standard Contractual Provisions*, 66 U. CHI. L. REV. 821 (1999).

42. Charny, *supra* note 41, at 843-44.

distorting effects of legal enforcement are emphasized. Perhaps ironically, autonomous conceptualism, divorced from life but not wholly alien to it if my conjectures about its intuitionist base are entertained, is a better candidate for protecting life from law. At least this may be so if formalist law is limited in ways that leave empty spaces for life. I postpone the question whether this is possible for a moment.

Let me address, briefly, one last criticism of autonomous conceptualism not yet noted. It is that formalism is impractical in a complex, heterogeneous and dynamic society. This claim is typically made with respect to the United States and is therefore typically accompanied by a concession that formalism operates, perhaps successfully, elsewhere.[43] I have three responses to these lines of argument.[44]

First, while it is surely the case that change occurs and may require change in law, the issue of change is far more important in an anti-formalist, purposive and instrumentalist conception of law than within a formalist conception. Law, in the former, is an instrument of planning on the assumption that law pervasively directs activity. Law, conceived as having this degree of responsibility for society is easily viewed as necessarily dynamic in a dynamic society. This, however, is not the role of law in the formalist conception, or, at least, in the formalist conception I wish to defend. If society operates, if not quite independently of law, at least independently of particularized direction by law, social change does not imply an urgent need for legal change.

Second, what is often meant by change is not change in fundamental social conditions or in technology, but change in intellectual fashion. Thus, the move from a formalist common law to social engineering in the progressive and New Deal eras was predicated in part on the idea that social conditions had changed, requiring new and different law. Yet it has become apparent that large aspects of this new and different law were substantial mistakes, requiring the dismantling of much of the legislation generated in these eras.[45]

Finally, when anti-formalists invoke the facts of complexity against formalism they assume that the proper response to these phenomena is to manage them. This is not surprising, it reflects a rationalist bias to the effect that greater complexity requires greater measures of control in service of articulated objectives. There is, however, an alternative response to complexity. It is that complexity requires less, not more managerial direction. Passivity in the form of complexity is counterintuitive to the rationalist, but it is obviously supportable

---

43. *E.g.*, POSNER, *supra* note 22, at 264-65.

44. I rely in what follows largely on Richard Epstein, *The Static Conception of the Common Law*, 9 J. LEGAL STUD. 253 (1980).

45. *E.g.*, POSNER, OVERCOMING LAW, *supra* note 5 at 220-21. *Cf.* CASS R. SUNSTEIN, AFTER THE RIGHTS REVOLUTION, RE-CONCEIVING THE REGULATORY STATE ch. 3 (1990) (recounting regulatory failure from pro-regulatory perspective). Critiques of Progressive Era, New Deal and Post-New Deal regulation are of course legion. *See* THE REGULATED ECONOMY: AN HISTORICAL APPROACH TO POLITICAL ECONOMY (Claudia Goldin & Gary Libecap eds., 1994); George Stigler, *The Theory of Economic Regulation*, 2 BELL J. ECONOMICS 3 (1971).

both by reference to theories of spontaneous order and by evidence in experience that attempted management of complexity fails.[46]

II. FORMALISM AS RULES

Another understanding of formalism is that the law consists, or should consist of rules.[47] The standard argument favoring rules rests upon an appeal to rule of law values: Rules enable those subject to them to predict the legal effect of their behavior and therefore enable coordination; rules preclude discretion and enable a claim that we are governed by law, not men; rules ensure that law is prospective, not retroactive.[48]

Rules should be distinguished from principles, standards, or rules of thumb in that rules direct particular legal conclusions or are more determinate than these alternatives. This implies strict application: the judge or other legal actor committed to rules is not free to make a decision on the basis of what seems best under the circumstances, nor is she free to ignore the rule where following the rule would produce a result she deems absurd, nor is she free to base her decision on the rule's purpose where the rule's directive in the circumstances of the case seems to her inconsistent with that purpose.[49]

Recall that formalism, understood as an autonomy claim, is non- or anti-instrumental, so it may be understood as rejecting the idea that law should be applied so as to achieve its purposes. This may seem odd. Most, if not all legal rules can be assigned plausible, functional purposes, and many can be plausibly said to serve such purposes. It is nevertheless obviously possible to seek to apply such rules in particular cases without reference to such purposes. A strong version of a rule utilitarian perspective and rejection of an act utilitarian perspective suggests as much.[50]

An implication of devotion to rules is that a rule's addressee may with impunity circumvent the rule though strict compliance with it, as by engaging in the evil, or a substantially similar evil, targeted by a rule while nevertheless simultaneously adhering to the rule.[51] Formalism may be understood as a theory of law that tolerates this activity. Thus, the form behavior takes, not the substantive nature of the behavior or the consequences of the behavior, is, for the formalist, controlling.[52] Indeed, a prominent feature of classical formalism was that its adherents openly advocated adherence to principle and rule even where

---

46. *E.g.*, HAYEK, *supra* note 32, at 119-208; MICHAEL OAKESHOTT, RATIONALISM IN POLITICS 5-42 (1962).

47. *E.g.*, Larry Alexander, *"With Me, It's All er Nuthin": Formalism in Law and Morality*, 66 U. CHI. L. REV. 530 (1999); Schauer, *supra* note 6.

48. *E.g.*, Antonin Scalia, *The Rule of Law as a Law of Rules*, 56 U. CHI. L. REV. 1175 (1989).

49. *See* FREDERICK SCHAUER, PLAYING BY THE RULES, A PHILOSOPHICAL EXAMINATION OF RULE-BASED DECISION-MAKING IN LAW AND IN LIFE 96-100 (1991).

50. *See* John Rawls, *Two Concepts of Rules*, 64 PHIL. REV. 3 (1955).

51. The doctrine of independent legal significance in corporate law is an example. *See* Hariton v. Arco Electronics, Inc., 182 A.2d 22 (Del. Ch. 1962), *aff'd*, 188 A.2d 123 (Del. 1963).

52. *See* KATZ, *supra* note 37.

they conceded that the result would be unjust, unfair or absurd.[53] This harsh notion is traceable to the very nature of the idea that the law consists of rules and compliance with law consists of following rules. If rules are suspended when they generate absurd results, they are no longer rules.[54]

Formalist rule worship may also be understood as entailing a theory of adjudication, specifically, "mechanical adjudication."[55] The theory is that rules may be applied to facts mechanically: rules reference sets of facts, so when the relevant set appears, the rule is applied and when it does not the rule is not applied. This conception is of course often attributed to lay persons and to entering law students, and when so attributed is always accompanied by the view that is hopelessly naive. It is, of course, often also attributed by judges to themselves; judges often justify their decisions on the basis that rules compel those decisions.

The formalist adjudicative theory thus depicted entails a deductive procedure. It is deductive in the sense that a rule as a major premise and a set of facts as a minor premise generates a right answer. A formalist legal opinion is one, then, that justifies the result reached by employing a syllogism of this type.

The standard critiques of formalist rule worship may be divided into two basic categories. First, rules have substantial defects.[56] As they are inevitably over- and under- inclusive, they fail to achieve their purposes where these purposes would be furthered by applying the rule to circumstances that the rule's language does not reach or would be furthered by not applying the rule in circumstances the rule's language does reach. Rules can produce absurd results in some circumstances. Absurd, that is, in that some value or norm would be violated by application of the rule, or some desired result would not be reached if the rule were applied. Rules suppress facts by rendering only some facts relevant to the rule, while facts left out by the rule are, by virtue of values, objectives or expectations, important. Anti-formalists will therefore think it desirable that judges refuse to apply rules or to stretch rules to serve their purposes, that they decline to apply rules where application produces absurd results, and that they formulate standards, rather than rules. Standards enable contextualized assessment and judgment, taking into account more facts and circumstances, and permit direct application of purpose and principle without the mediation of a rule.[57]

---

53. CHRISTOPHER COLUMBUS LANGDELL, SUMMARY OF THE LAW OF CONTRACTS 20-21 (1880), *quoted in* Grey, *supra* note 10 at 3, 15.

54. *E.g.*, Alexander; *supra* note 47, at 531, 547, 553-55; SCHAUER, *supra* note 49, at 116.

55. *Cf.* Roscoe Pound, *Mechanical Jurisprudence*, 8 COLUM. L. REV. 605 (1908) (objecting to what I have here termed autonomous conceptualism).

56. *See, e.g.*, POSNER, *supra* note 2, at 44-49; SCHAUER, *supra* note 49, at 100-02; SUNSTEIN, *supra* note 5, at 121-35.

57. POSNER, *supra* note 2, at 44-49; *cf.*, SUNSTEIN, *supra* note 5, at 136-47 (balancing "factors" as alternative to rules). On the rules versus standards debate generally, see, for example, Alexander Alienikoff, *Constitutional Law in the Age of Balancing*, 96 YALE L. J. 943 (1987); John Calfee & Richard Craswell, *Some Effects of Uncertainty on Compliance With Legal Standards*, 70

The second basic critique is that adjudication by reference to rule—the mechanical adjudication generally attributed to classical formalism—is highly implausible.[58] Adjudication as syllogism, with the rule as major premise and facts as minor premise may be that which is expressed in a formalist decision, but this expression covers up the hard and problematic work that goes into generating these premises. Rules cannot themselves be identified through deduction, for there can be multiple and conflicting rules plausibly invocable. A choice of rule is therefore necessary, and the formalist who relies simply on syllogism has failed to justify his choice. There are gaps among and between rules, so the formalist who pretends to apply a prior rule to the gap has failed to justify what is in effect a new rule. Rules, particularly the legislature's rules we call statutes, often employ words with no clear referents, so the formalist who insists, for example, that the words "manufactured goods" apply, by virtue of the meaning of these words, to the fact of an "eviscerated chicken"[59] has again failed to justify his decision.[60]

These failures of justification are failures of formalist adjudication: the constrained, mechanical, or deductive technique attributed to formalism *cannot* work. We may add to these problems the questionable character of facts and of factual findings.[61] Our means of resolving factual disputes are weak and often distorted both by our processes and by human frailties. The facts we find, even absent dispute, are at best partial under a rule regime; much that is arguably relevant is left out. The anecdotal facts of particular disputes are not the systematic facts necessary to formulating social policy, even if expressed in rules.

VA. L. REV. 965 (1984); Richard Epstein, *The Risks of Risk/Utility*, 48 OHIO ST. L.J. 469 (1987); Jason Scott Johnston, *Uncertainty, Chaos, and The Torts Process: An Economic Analysis of Legal Form*, 76 CORNELL L. REV. 341 (1991); Louis Kaplow, *Rules Versus Standards: An Economic Analysis*, 42 DUKE L.J. 557 (1992); Duncan Kennedy, *Form and Substance in Private Law Adjudication*, 89 HARV. L. REV. 1685 (1976); Pierre Schlag, *Rules and Standards*, 33 UCLA L. Rev. 379 (1985);.

58. BENJAMIN CARDOZO, THE NATURE OF THE JUDICIAL PROCESS 112-15 (1921); POSNER, *supra* note 2, at 42-61; Felix Cohen, *The Ethical Basis of Legal Criticism*, 41 YALE L.J. 201, 215-19 (1931); Duncan Kennedy, *Legal Formality*, 2 J. LEGAL STUD. 351 (1973); Joseph Singer, *The Player in the Cards: Nihilism and Legal Theory*, 94 YALE L.J. 509 (1988).

59. *Cf.* Interstate Commerce Comm'n v. Krobin, 113 F. Supp. 599 (N.D. Iowa 1953), *aff'd*, 212 F.2d 553 (8th Cir. 1954) (presenting these facts and issue, but not necessarily displaying this reasoning).

60. Michael Moore, *The Semantics of Judging*, 54 S. CAL. L. REV. 151 (1981). Another argument is that words have no core, linguistic meanings. *E.g.*, James Boyle, *The Politics of Reason: Critical Legal Theory and Local Social Thought*, 133 U. PA. L. REV. 685 (1985); Lon Fuller, *Positivism and Fidelity to Law—A Reply to Professor Hart*, 71 HARV. L. REV. 630 (1958);. The argument has been demolished by Professor Schauer. SCHAUER, *supra* note 49, at 55-61.

61. JEROME FRANK, COURTS ON TRIAL 316-21 (1949).

A. Formalist Adjudication

What may be said in response to these critiques? Let me begin in reverse order by addressing the problem of formalist adjudication, understood as the unproblematic application of rules to facts. It will turn out that problems of adjudication are related to the critique of rules, as such, so my discussion of adjudication will lead to discussion of that critique.

A typical and, I think, persuasive response to the critique from the impossibility of unproblematic application is some version of a hard case/easy case dichotomy.[62] The defense focuses upon the easy case and observes that in fact rules, including legal rules, are unproblematically applied to facts all the time. Without contending that meaning resides in language or that facts are easily identified, most cases are resolved before they ever enter the realm of formal adjudication because in most cases there is agreement about the meaning of the rule, the facts and the application of rule to facts. It is the hard case that is adjudicated, or it is the hard case that attracts an appeal and is the subject of interest. It is, therefore, only the hard case that displays the problems emphasized by the critiques.

On this account, formalist "adjudication" works most of the time. In particular, it works in the hands of layman and lawyers outside of court when engaged in the activity of law compliance or of Holmesian prediction of what judges will do "in fact." Realist critiques of formalist adjudication thus betray legal realism's peculiar focus upon, indeed fixation with the judge.

What, however, of the hard case? It seems apparent to me that the critique of formalist adjudication clearly works in *some* hard cases. In particular, it works where there is no plausibly applicable rule available to resolve a case, where two plausibly applicable rules conflict, and where the rule in question has no clear referents.[63] Adjudication in these cases is indeed problematic. A "grab bag" of techniques, perhaps best described in terms of "practical reason" must be invoked to resolve the hard case, and the formalist description of adjudication is an inaccurate depiction of the grab bag.[64] But this assumes that it is formalist adjudication, in the sense of unproblematic application of rule to fact, that is being assessed. What of a formalist recommendation that hard cases be resolved so as to become easy cases in the future?

There is nothing in the critique of formalist adjudication that would preclude such a recommendation. Thus, the formalist confronted with a hard case of the type indicated may resolve it by establishing a rule (not a standard), by seeking to employ words with clear referents in stating the rule, and by minimizing the

---

62. *E.g.*, H.L.A. HART, THE CONCEPT OF LAW 122-32 (1961); DUNCAN KENNEDY, A CRITIQUE OF ADJUDICATION 159, 275 (1997); SCHAUER, *supra* note 49, at 192-59; SUNSTEIN, *supra* note 5, at 128-29; H.L.A. Hart, *Positivism and the Separation of Law and Morals*, 71 HARV. L. REV. 593 (1958); Frederick Schauer, *Easy Cases*, 58 S. CAL. L. REV. 399 (1985).

63. In my view, "plausibly applicable" means most locally applicable. *See* SCHAUER, *supra* note 49, at 188-91. Thus, the case contemplated is one of conflicting local rules, not one of arguable "conflict" between a local rule and a more abstract or distant one.

64. POSNER, *supra* note 2, at 73.

set of facts that will be deemed relevant under the rule. The primary criterion for resolving the hard case therefore becomes "formulate that resolution that will best enable formalist adjudication in the future."[65] There are, of course, institutional constraints on the ability of the formalist to do these things. A common law judge is no doubt less able to do so than a positivist's sovereign. But it remains the case that formalist adjudication can be understood as prospective and programmatic as a conscious effort to turn today's hard case into tomorrow's easy case.[66]

There is another category of case said to be "hard" that formalists will not regard as hard in the same sense. This is the category of the absurd result or of application of the rule not serving its purpose or of the inapplicability of the terms of the rule permitting the evil targeted by the rule. What is hard about such cases is not a matter of the rule's apparent meaning. It is perfectly clear that the rule means what it says in the context of the facts presented. It is perfectly clear precisely because it would otherwise make no sense to claim that this meaning produces an absurd result or fails to serve its purpose.[67] These cases are hard not because of a question of meaning, but because of a normative issue: should the decision maker tolerate absurd results or results inconsistent with purpose?

I think most law professors and many judges would answer "no" to this question. Indeed, one is warranted in saying that contemporary law generally reflects this answer. I also think, however, that there are very good reasons for an affirmative answer. These reasons have largely been supplied by others,[68] so I will merely summarize some of their points and add a word.

The basic thrust of the defense of formalist adjudication in hard moral cases is that departures from the known meaning of a rule in such a case undermine, or destroy the reasons for rules. These reasons, interestingly, are *consequentialist* reasons; they supply good utilitarian (in a broad sense) grounds for preferring rules over standards or good instrumental reasons for "ruleness." Notice then, that a defense of what I have been calling formalist adjudication leads to a defense of rules.

B. Rules' Function

Consider in particular the following, highly simplified summary of Professor Larry Alexander's consequentialist defense of rules:[69] (1) people face coordination problems (they need to know how others will act and what to do in the case of disagreement), (2) rules solve this coordination problem by supplying "authoritative settlements" and do so in ways superior to particularized authoritative direction in each case of questioning what to do because (3) the

---

65. *See* Scalia, *supra* note 48, at 1183-87.
66. This, indeed, was Justice Holmes' program. *See* Grey, *supra* note 10, at 44.
67. SCHAUER, *supra* note 48, at 55-62, 213-15.
68. *See id.* at 158-66; Alexander, *supra* note 47.
69. Alexander, *supra* note 47.

costs of more particularized modes of authoritative settlement are prohibitive.[70] There are, of course, some necessary caveats. Rules, to serve their function must be determinate in meaning (indeed Professor Alexander defines "rule" by reference to this quality) and must be knowable.[71] They should therefore usually be general and few rather than specific and many (as complexity undermines knowability).[72] Notice that rules are not in Alexander's (and for that matter, F. A. Hayek's similar) depiction a solution to the problem of "bad men," persons not motivated to do the right thing. Rather, they are solutions to the problem of ignorance knowing what the right thing to do is.[73]

One alternative to rules, and a form of particularized authoritative settlement, is "standards." The usual example of a standard, although there are reasons to think it a bad example, is negligence failing to exercise the care a reasonable person would exercise under the circumstances.[74] Standards may be distinguished from rules on the basis that rules are determinate and standards are not. The difficulty with standards, in Professor Alexander's analysis, is that they duplicate the problems rules are supposed to solve. That is, as standards are indeterminate, there will be disagreement about their meaning in particular cases; they will fail to inform us of what to do. This is not always so. A reasonable person standard is determinate (and therefore a rule) if everyone or nearly everyone in a community agrees about what a reasonable person should do. But the uncertainty and disagreement that the law is to minimalize are usually merely duplicated in standards.

If this is so, it should be clear why application by reference to the underlying "purpose" of a rule or refusal to apply a rule where doing so produces absurd results is "wrong" and strict adherence to rules is "correct" from the formalist perspective: these non- or anti-formalist actions turn rules into standards.[75] Adjudication by reference to purpose in preference to known plain meaning resurrects controversy over purpose, particularly given the possibility of ascending abstraction in characterizing purpose.[76] Avoidance of absurd result assumes agreement about absurdity, but there is very often no such agreement.

Perhaps, however, this equating of purpose-oriented interpretation and absurd result avoidance with substituting standards for rules is too extreme. If a standard can be a rule where everyone agrees about its meaning in context, then

---

70. *Id.* at 531-40.

71. *Id.* at 542-45.

72. *Id.* at 545.

73. *Id.* at 549. Hayek's positions were derived from a general interest in the problem of ignorance; he, unlike most economists, largely ignored problems of self-interest. *See* Marina Bianchi, *Hayek's Spontaneous Order, The "Correct" Versus the "Corrigible" Society, in* F.A. HAYEK, COORDINATION AND EVOLUTION 232-51 (Jack Birner & Rudy Van Zijp eds., 1994).

74. *E.g.*, POSNER, *supra* note 2, at 44. For the argument that it is a bad example, see *infra* note 128.

75. Alexander, *supra* note 47, at 547.

76. On problems with purpose, see Frank Easterbrook, *Statutes Domains,* 50 U. CHI. L. REV. 533 (1983); Max Radin, *Statutory Interpretation*, 43 HARV. L. REV. 863, 876-77 (1930).

it ought to be possible for a similar agreement to occur with respect to purpose and absurdity. Perhaps, but the problem is that of the slippery slope.[77] A legal practice in which purpose and absurdity permit departures from plain meaning in cases of such agreement will lead to one in which such departures are routinely made in cases of substantial and widespread disagreement. This, indeed, happens often in our contemporary practice.[78]

I wish to add to this summary of a defense of rules an observation about the function of law it assumes. I do so because this function may tell us something about formalism apart from its preference for hard rules. The function contemplated is coordination of action in the face of uncertainty. That is a sufficiently broad statement to encompass numerous versions of "coordination," but I wish to narrow the notion of coordination in a way that renders it close to the assumptions and understandings of the classical formalists. The picture I wish to invoke is one in which persons are acting in service of their own ends and require law only for the purpose of not bumping into each other while doing so, or for the purpose of ensuring efficacy of exchange.[79] Once a rule is provided, compliance follows and *the law is left behind.* An interesting feature of this picture is that it further explains hostility to standards (and to other *ad hoc* modes of "authoritative settlement"). Specifically, the trouble with standards is that their uncertainties compel persons who otherwise would prefer to get on with their lives and leave the law behind them to engage in argument and participate in a process of public justification. This, of course, is why left-communitarians tend to be critical of rules and favor standards. It is, of course, also why libertarians tend to favor rules.

I should nevertheless make it clear that rules, even general rules, will not themselves implement a libertarian program. Hayek, at least at one point in his intellectual odyssey, thought that such rules would do the trick,[80] but he was, I think, wrong. The reason is that the substantive content, number and complexity of rules must be taken into account. It is quite possible for rules satisfying formal requisites to nevertheless so constrain the "negative liberty" Hayek advocated as to defeat his political program.[81] Consider, for example, that much of the law of the "administrative state" is comprised of inflexible command and control directives issued by administrative agencies in the form of regulations. These often produce absurd results,[82] and formalism as mindless rule worship is surely

---

77. *See* Frederick Schauer, *Slippery Slopes*, 99 HARV. L. REV. 361 (1985).

78. My candidate for a prime example of this phenomenon is *United Steelworkers v. Weber*, 443 U.S. 193 (1979).

79. The latter purpose may address problems of "cooperation" as well as problems of "coordination," but formalism's non-facilitative implications, *supra* note 10, may often result in non-cooperation.

80. HAYEK, CONSTITUTION OF LIBERTY, *supra* note 4, at 205-19.

81. *See* SUNSTEIN, *supra* note 5, at 156-61 (comparing mandatory, end-state directive rules versus privately adaptable rules).

82. *E.g.*, CASS SUNSTEIN, FREE MARKETS AND SOCIAL JUSTICE 271-97 (1997). *See supra* note 45.

a standard characterization of the law generated by bureaucracy.

Does this mean that I have given up on a defense of formalism—that I have conceded that it is the substance, not the form of the law's authoritative settlement that is important? I do not believe so. To say that substance matters is not to say that form does not. Rules have the tendencies depicted in the picture of persons leaving law behind and standards have the tendencies depicted in the picture of persons forced to engage in public justification. If the leaving law behind picture is attractive, as it is to me, rule preference is *an aspect* of the legal program that serves this picture.

C. Rules and Facts

Before leaving the matter of rules, I want to briefly pick up a theme about facts that I have thus far largely ignored. I suggested above that formalism may also be criticized for its uncritical reliance upon the "facts" found in legal proceedings.

It is not, however, clear that difficulties in establishing facts present a threat to formalism as "mechanical adjudication." There may well be factual uncertainty, but the formalist syllogism treats the minor premise as an assumption or stipulation. However messy factual determinations might be, the logical exercise proceeds after these determinations are made. It may, therefore, be possible to be both a formalist and a fact skeptic.

It has been said that classical formalists preferred "readily ascertainable facts."[83] They may be said, then, to have been indeed fact skeptics in the sense that they distrusted discretion in fact finding: The fewer the factual assumptions necessary to form minor premises the better. So, for example, objective rules were preferred to vague standards, as standards require or permit assessment of more facts. It might, therefore, be said that formalists ignore or de-emphasize facts in service of conceptual order. The complexities of human behavior and the multiple potential considerations arising from these complexities are threats to rules, so formalists suppress these complexities and considerations by giving primacy to rules.

Moreover, formalists are thought to prefer abstract and general rules over particularized or specialized rules. They prefer, for example, one law of contract, not multiple laws for distinct types of contracts or distinct contractual settings.[84] This entails suppression of factual difference through an assumption of greater homogeneity than may exist in fact. This suppression of factual difference also facilitates, however, the formalist aspiration to a complete, coherent system from which correct answers may be derived. It enhances the prospects for consistency where consistency is to be obtained at the levels of conceptual principle and rule rather than through particularized factual distinctions.

An insistence upon expanding the scope of factual inquiry goes hand in hand

---

83. Grey, *supra* note 10, at 11. *Cf.* Andrew Krull, *The Simplification of Private Law*, 51 J. LEGAL EDUC. 284 (2001) (contending that there is a contemporary tendency to simplify private law by rejecting fact-sensitive equitable inquiries).

84. GILMORE, *supra* note 1, at 82-83.

with standards, balancing tests and factor analysis, for the underlying notion is that judgment is to be made all things considered.[85] This, however, is precisely what formalism's emphasis upon rules condemns, for the reasons noticed above.[86] By contrast, formalism's suppression of facts goes hand in hand with formalism's distance from life and facilitates that distance. Notice that this is not a criticism of formalism; formalism's defense of its distance from life is consistent with its hostility to particularized decisions under standards and, therefore, its suppression of facts. Fact suppression limits law's intrusion into life, rendering the facts it suppresses nevertheless available for human judgment within the framework supplied by formalist rules.[87]

This, I think, is an answer to the common claim that the rigidity of rules and the suppression of facts by rules are alien to human judgment, or, at least, to preferred conceptions of human judgment. If the sociologists and institutionalists are correct, human behavior is largely scripted, a matter of rule following even outside law. Nevertheless, a more flattering picture of human choice, or, at least, of wise human choice, entails "all things considered" judgment. So, from the perspective of this picture, judicial (or other governmental) decision by inflexible reference to rules is denigrated, as by claiming that judges are not or should not be mere rule followers.[88] I, too, prefer the picture of wise judgment, all things considered, but it is not necessary to this ideal that it be the judge or other governmental functionary who exhibits wise judgment. The point of a rule (or, more accurately, of rules with a particular substantive orientation) is that it provides a framework within which such judgment may be exercised by persons other than governmental functionaries. It confers, in effect, the jurisdiction to be wise.[89]

Another criticism of formalist facts is that they are anecdotal—they fail to provide adequate data about systematic human tendencies. This, of course, is a pragmatic instrumentalist complaint: If law is conceived to be an instrument of comprehensive planning to service collectively determined ends, "legislative facts" are needed. It is, of course, also a complaint about common law adjudication generally, not just formalist adjudication (unless formalism is

---

85. *E.g.*, POSNER, *supra* note 2, at 44-49; SUNSTEIN, *supra* note 5, at 136-47.

86. *See supra* notes 70-81 and accompanying text.

87. *See supra* notes 42-43 and accompanying text, *infra* notes 107-19 and accompanying text.

88. POSNER, OVERCOMING LAW, *supra* note 5, at 489-92.

89. *Cf.* SCHAUER, *supra* note 49, at 158-66 (stating that primary function of rule is allocation of decision making authority). Notice, however, that this is potentially so in two senses. A rule can be viewed, as Schauer largely does, as retaining the authority to be wise (or foolish) in the original rule maker. It might also he thought, however, to confer the authority to be wise (or foolish) on persons subject to the rule. This latter sense may seem doubtful if one contemplates a directive rule. Consider, however, a rule requiring consideration for the legal enforceability of a promise. The maker of a promise has, under such a rule, the "discretion" to obtain legal enforceability through a demand for consideration and the discretion to perform, or not, if he fails to make this demand. Consider, also, a prohibition of theft. The prohibition withdraws the discretion of those subject to it to steal, but also confers the discretion (and possibility) of contracting for property transfer.

defined as common law adjudication). The complaint serves, for example, to justify displacement of common law adjudication by regulation through the supposed expertise of administrative agencies.[90] Whether or not administrative regulation in fact exhibits expertise in either the identification of systematic facts or in their assessment, the important point for present purposes is to recognize that the function of law is quite distinct in the "administrative state" from that proposed above as an explanation of formalist rules and of formalist suppression of fact.[91] The function envisioned for formalist law, recall, was a matter of limited coordination. The function envisioned by the administrative state is comprehensive, top-down planning in service of collectively determined ends. There is obviously a greater need for facts in the latter than the former.

III. FORMALISM AS EMPTY SPACES

A prominent feature of legal realism, and, later, of critical legal studies, is a rejection of the idea of the empty space—an area in which persons are free from law. Actually, there appear to be two distinct but related realist ideas here. First, there is the Hohfeldian idea that liberty (in Hohfeld's terminology "privilege") is distinct from legal right.[92] Thus, the law does not in many instances preclude interference by others with liberty; persons in those instances may harm others with legal impunity. When the law does intervene, when it recognizes a right, it simultaneously imposes a duty, so one person's right is merely the legal enforcement, or threat of enforcement, of another person's duty. One upshot of Hohfeldian analysis is the recognition that legal rights are constraints on liberty. Another is that concepts like property refer to bundles of legal relationships, not to real things in the world. Still another is that one cannot suppose, as classical formalists are said to have done, that, because the law recognizes a liberty to do *X*, in the sense that the law permits *X*, that there is a right to do *X*, in the sense that the law will impose a duty not to interfere with one's doing of *X*.[93]

An implication of this last point is that classical formalists were wrong in supposing that rights could be logically derived from privileges.[94] Another is that the Millian concept of liberty as the freedom to pursue one's own ends so long as one does not harm others is not a viable explanation of the legal system given the extent to which that system privileges the infliction of harm.[95] This, in turn, implied that no single principle could explain when the law would and would not intervene to prevent harm, a substantial threat to classical formalism's conceptualistic, deductive system.[96]

---

90. *See* SUNSTEIN, *supra* note 45.

91. *See* Gjerdingen, *supra* note 33; Mashaw, *supra* note 33.

92. Hohfeld, *supra* note 16.

93. *See generally* Joseph W. Singer, *The Legal Rights Debate in Analytical Jurisprudence from Bentham to Hohfeld*, 1982 WIS. L. REV. 975.

94. *Id.* at 997-98.

95. *Id.* at 1022.

96. Some legal economists believe, of course, that the principle of efficiency, here in the

Nevertheless, Hohfeldian privilege or liberty seems clearly to recognize empty spaces in the law: areas of freedom from law, or, in effect, states of nature.[97] How, then, can I claim that realism rejected the idea of an empty space? The answer lies in a further aspect of Hohfeld's thought, one emphasized, in particular, by the legal realist Robert Hale.

A response to Hohfeld was that the realm of liberty (privilege) was outside law, not a part of it. If the law recognizes no duties within the empty space of privilege, then that space is empty of law.[98] To this Hohfeld replied that "[a] rule of law that *permits* is just as real as a rule of law that *forbids* . . . ."[99] Thus, a judge who finds for a defendant on the basis that the defendant had no duty of noninterference has made a legal decision. How far might this characterization be pushed? Hale pushed it to rather extreme lengths: Not only is the decision to deny a legal duty a legal decision, it is a delegation of state power to the defendant holder of Hohfeldian privilege.[100] Since liberty is recognized by law, the acts undertaken within it are state acts. Indeed, Hale saw state-based coercion everywhere: A voluntary contractual exchange was, for Hale, "coerced" by the fact that both parties are legally entitled to withhold consent.[101] Hale's thought is evident in the oft-repeated contemporary view that any given "private" preference, realm or decision is in fact legally constructed by virtue of a background of state determined entitlements and is therefore "really" a "public" preference, realm or decision.[102]

So realism, and much contemporary thought, rejects the empty space idea, not in the sense that it fails to recognize liberty to harm others in the law, but, rather, in the sense that it denies that this liberty is apart from law. The realist claims are, then, that law permeates liberty, that there is no private realm, and that the private is publicly constructed.

What has all this to do with formalism? If formalism is that which its critics'

---

guise of pecuniary versus non-pecuniary externality, explains at least the common law.

97. Duncan Kennedy & Frank Michaelman, *Are Contract and Property Efficient?*, 8 HOFSTRA L. REV. 711, 715, 727-28, 754 (1980). While states of nature (or pockets thereof) are extreme examples of empty spaces, it should be noted that I have a broader idea in mind. *See infra* note 111. Thus, in my scheme, there can be an "empty space" generated by legally enforced property entitlements and contract rules even though these entitlements and rules obviously presuppose a state. So a "state of nature" in the pristine sense is not the intended meaning of my invocation of the phrase. A "state of nature" is, rather, a way of understanding Hohfeldian privilege, and such privileges may exist within a background set of entitlements and rules entailing Hohfeldian rights and duties.

98. J. AUSTIN, THE PROVIDENCE OF JURISPRUDENCE DETERMINED 290 n.* (1832).

99. Hohfeld, *supra* note 16, at 42 n.59.

100. *E.g.*, Robert Hale, *Coercion and Distribution in a Supposedly Non-Coercive State*, 38 POL. SCI. Q. 470 (1923); Robert Hale, *Force and the State: A Comparison of Political and Economic Compulsion*, 35 COL. L. REV. 149 (1935).

101. Robert Hale, *Bargaining, Duress, and Economic Liberty*, 43 COLUM. L. REV. 603 (1943).

102. *E.g.*, HORWITZ, *supra* note 2, at 193-212; CASS SUNSTEIN, THE PARTIAL CONSTITUTION, 68-92, 162-94 (1993).

attack,[103] the realist view is that formalists both fail to recognize that the law permits the infliction of harm and erroneously insist upon the existence of a realm of "liberty" apart from and ungoverned by law. Is this so? There is an obvious affinity between the empty space as liberty notion and my earlier claims that formalism seeks to leave law behind and to protect life *from* law. Moreover, the empty space idea fits, rather neatly, other features of formalism. The point, recall, of both an autonomous, conceptualistic basis for law and of rigid rules as expressions of law is to confine judicial discretion and to enhance stability and predictability. These objectives, if realized, would generate an undirected order within which individuals would pursue their individual projects.[104] Classical formalist commitments to "liberty" would then seem to follow from classical formalist conceptions of law. The realists attacked not merely the formalist commitment, but the very idea of liberty as a realm untouched by law.

Can the empty space idea be defended? One defense, ironically, is that critics of the empty space idea are themselves formalists.[105] To say that private action is "really" public action, or that the private is legally constructed and therefore "political" is to engage in absolutist conceptualism, for it is both true and not true that the private is private and that the private is public. It is true that persons are empowered to act within the private realm by virtue of a "baseline" set of background entitlements recognized in the traditional common law.[106] It is not true that this baseline either directs particular actions within this realm or, indeed, even addresses what particular actions will be undertaken within this realm.[107] More importantly, the fact of a baseline does not imply that it is itself consciously planned or constructed. Nor does recognition of the baseline justify

---

103. I recognize that formalism cannot simply be that which realists attack. I mean, instead, that which realists (etc.) attack *as formalism*, and I think it apparent that "empty spaces" are conceived by many critics of formalism as part and parcel of formalism. *E.g.*, DUXBURY, *supra* note 7, at 106-11; HORWITZ, *supra* note 2, at 155. *See* GILMORE, *supra* note 1, at 55 (Holmes' formalism greatly limited liability); SUNSTEIN, *supra* note 102, at 40-67, 112-19 (linking formalism as mechanical legal interpretation with substantive commitment to status quo distributions, and latter to *Lochner*). *But see, e.g.*, HOVENKAMP, *supra* note 35, at 174-75 (denying link between formalism and effort, as in *Lochner*, to constitutionalize common law version of liberty); SUNSTEIN, *supra* note 5, at 118-20 (denying association of Rule of Law with free markets). It is possible to separate formalist method from formalist normative commitment, but, as I suggest immediately below and *infra*, text and notes 167-71, I believe that there are in fact functional linkages between the two.

104. This, at least, was Hayek's vision. HAYEK, LAW, LEGISLATION AND LIBERTY, *supra* note 4, at 106-10, 118-22.

105. POSNER, OVERCOMING LAW, *supra* note 5, at 281-84. *Cf.* Larry Alexander, *The Public/Private Distinction and Constitutional Limits on Private Power*, 10 CONST. COMMENT. 361 (1993) (claiming that legal and conceptual breakdowns of public/private distinction have little normative force).

106. SUNSTEIN, *supra* note 102, at 40-92.

107. This is Hayek's reply to Hale. 2 F.A. HAYEK, LAW LEGISLATION AND LIBERTY: THE MIRAGE OF SOCIAL JUSTICE 37-38 (1976).

a program of conscious reconstruction. This charge, that the critics turn out to be formalists, is a highly attractive rhetorical point. Unfortunately, it is obviously not one upon which I can rely given a project of defending formalism.

So allow me to offer three defenses of the empty space idea distinct from defense through the charge of hypocrisy. The first may be termed a semantic defense. To say, with the realists, that withholding consent to a contract is "coercion" or that there is no "private" realm is to attempt the destruction of perfectly useful terms on the highly doubtful premise that persons who employ such terms are unaware of the legal nature of the institutional structure within which such perfectly useful terms are employed.[108] The formalist who denies that there is an implicit allocation of entitlement in the law's refusal to assess behavior would of course be mistaken, but no sophisticated formalist would deny this. Hayek certainly did not.[109] The empty space idea is precisely that the law's refusal to recognize an obligation confers power on persons and frees such persons from justifying their actions in terms of public ends. That the law, even contemporary law, in fact contains such empty spaces requires that the realist bent on denying the private and insisting on the ubiquity of state coercion must invent new, and often more obscure terms to describe these phenomena.

Second, the phenomena do in fact exist in the law; there are empty spaces. Consider two examples: (1) The business judgment rule generally precludes judicial assessment of corporate director decisions absent conflicts of interest and therefore leaves managerial decision making "unregulated," even though the corporation and the position of power of the board of directors within it are in important *senses* creatures of law.[110] (2) The employment at will doctrine precludes judicial assessment of an employer's decision to discharge an employee (and, for that matter, an employee's decision to resign) even though the very identification of who is an employer and who is an employee is a function of a set of background entitlements recognized and enforceable by law.[111]

---

108. Richard Epstein, *The Assault That Failed: The Progressive Critique of Laissez Faire*, 97 MICH. L. REV. 1697, 1700, 1704 (1997).

109. *See, e.g.*, HAYEK, *supra* note 32, at 112-16.

110. *See, e.g.*, Aronson v. Lewis, 473 A.2d 805 (Del. 1984); Shlensky v. Wrigley, 237 N.E.2d 776 (Ill. App. Ct. 1968); Charles Hansen, *The ALI Corporate Governance Project: Of The Duty of Due Care and the Business Judgment Rule, a Commentary*, 41 BUS. LAW. 1237 (1986).

111. *See generally* Andrew P. Morriss, *Bad Data, Bad Economics and Bad Policy: Time to Fire Wrongful Discharge Law*, 74 TEX. L. REV. 1901 (1996); Edward B. Rock & Michael L. Wachter, *The Enforceability of Norms and the Employment Relationship*, 144 U. PA. L. REV. 1913 (1996).

Let me clarify the notion of an empty space. There are, in my conception varieties and degrees of empty spaces; some spaces are more empty of law than others. For example, one device by which empty space may be created or expanded is that of constricting the realm of tort and expanding the realm of contract. The realm of contract is not an empty space in the same sense that a state of nature is an empty space; there are rights and duties within the space generated by contract. Nevertheless, the contractual space is "less full" of law than space governed by tort in the obvious senses that the rights and duties generated by contract find their source in the parties'

Third, it is a very good thing that there are empty spaces and it would be a significantly better thing if there were more and wider empty spaces. Leaving aside the many persuasive instrumental and consequentialist reasons for such empty spaces as those created by the business judgment rule and the employment at will rule, let me offer a reason for the goodness of empty spaces more in keeping with what I am characterizing as a formalist stance. I said above that the point of the empty space was freedom from public justification. It may be

---

agreement to these, not in an externally imposed direction. I am aware that the realm of contract can be characterized as full of directive law. *E.g.*, Jean Braucher, *Contract Versus Contractarianism: The Regulatory Role of Contract Law*, 47 WASH. & LEE L. REV. 697 (1990). I do not share that view. *See* Barnett, *supra* note 38. Within the realm of contract, a further means of expanding empty space is that of expanding the realm of default terms and limiting, or eliminating, the realm of mandatory terms. Within the realm of remedies, the device is that of favoring those that force market transactions, such as specific performance and injunction, and disfavoring those that entail judicial assessments, such as damages.

Now, one theme that runs through these examples is a program of withdrawal from mandatory and directive law, so an empty space is by reference to this theme freedom from and freedom to contract. Another theme, however, is limiting occasions for judicial assessment, and this theme will not only entail a withdrawal from directive law, *it will entail a withdrawal from facilitative law.* It will entail, for example, limitations on freedom *to* contract, because it implies a reluctance to engage in problematic factual assessments. *See, e.g.*, Alan Schwartz, *The Default Rule Paradigm and the Limits of Contract Law*, 3 S CAL. INTERDISC. L. REV. 389 (1993); *cf.* GILMORE, *supra* note 30, at 52-54 (contradiction between bargain theory and absolute liability potentially resolved by desire to limit litigation).

An example is the strict bargain principle of contract, a principle that excludes firm offers from enforcement and therefore fails to facilitate exchange. *See supra* notes 25-32 and accompanying text. Another example entails rejecting the notion that courts are capable of identifying the "reasonable expectations" of shareholders in closely held corporations, *e.g.*, Robert B. Thompson, *Corporate Dissolution and Share-Holders' Reasonable Expectations*, 66 WASH. U. L.Q. 193 (1988). That notion seems to me, highly doubtful if the inquiry is understood as empirical. If, instead, the inquiry is understood as imposing tort-like mandatory terms, it is directive and therefore suspect from the perspective suggested here. But it is at least arguable that withdrawing from reasonable expectations inquiries will deter initial investments. A final example: At one point in the history of corporate law an interested director contract was simply voidable; later, such a contract became enforceable if "fair." The earlier rule is a formalist rule if formalist rules are, as I advocate, designed to limit judicial assessment. The later rule requires inquiring into the open-ended matter of fairness, and risks the imposition of conception of fairness alien to the understandings of parties to the corporate "contract." Nevertheless it enables mutually beneficial deals precluded under the earlier rule.

The point is that formalist non-direction and formalist non-assessment will necessarily entail the withdrawal of law from the enterprise of facilitating exchange and therefore relegating that project to aspects of society outside law. In law and economics lingo, the formalist project of expanding empty spaces operates, in effect, as a counterfactual but strong presumption of zero transaction costs and as a more factually supportable assumption of extremely high administrative costs.

true—and certainly would be under a pragmatic instrumentalist regime—that the empty space as a class or category of conduct may be assigned a "public justification," the justification, for example, of maximizing social wealth. But it remains the case that, once recognized, the empty space is a haven *from* public justification—an area within which one may leave law behind. It seems to me that the goodness of this notion, from the point of view of an individualist tradition, is self evident. It is reflected, in highly imperfect forms, in post-New Deal constitutional law,[112] albeit not within so-called economic realms. And it is reflected, again imperfectly, within these realms in the doctrinal examples I have given. I will not seek to defend an individualist tradition here, but I do wish to make clear what I take to be the nature of the goodness of the empty space claimed by that tradition. It is precisely that articulate justification for (formally private) choice is not asked, let alone required.

My final defense of empty spaces rests on the agenda of the critics of those spaces. The agenda, I claim, is precisely a denial of the goodness of the empty space postulated by the individualist tradition. The critics, it must be recognized, come from both ends of the political spectrum, but allow me to concentrate upon what I take to be the legal realist tradition. Realism's denial of the empty space is premised, I submit, upon a pervasive, indeed organic conception of law in both descriptive and normative senses. The descriptive prong of this conception, we have already encountered: there is no such thing as a private realm because each choice within the realm is traceable to a legal allocation of power. The normative prong goes like this: as the private realm does not exist, it is not an obstacle to a centralized, instrumental and purposive collective assessment, which assessment is itself a good thing.[113] The goodness of such an assessment, from this realist perspective, is precisely that articulate justification of formerly private choice is to be required.[114]

It might be thought that I exaggerate, but I think I do not. When it is said, as it sometimes is currently said, that we have too much law, when, for example, Professor Gilmore's notion that "[i]n hell there will be nothing but law"[115] is quoted, the speaker is recognizing, in my terminology, the contraction of empty spaces. This phenomenon of contraction is evident, for example, in contemporary threats to the continued viability of my examples of empty spaces, the business judgment rule and the employment at will rule. It is a phenomenon, however, I think pervasive. I suspect that for every example of a common law empty space, particularly where the space is generated by a hard looking legal rule, one may find either progressive retreat from the rule or the parallel development of an alternative body of law that undermines the empty space

---

112. Louis Michael Seidman, *Public Principle and Private Choice: The Uneasy Case for a Boundary Maintenance Theory of Constitutional Law*, 96 YALE L.J. 1006 (1987).

113. *See, e.g.*, SUNSTEIN, *supra* note 45, at 160-92; SUNSTEIN, *supra* note 102, at 40-92.

114. *Cf.* Louis Michael Seidman, *The State Action Paradox*, 10 CONST. COMMENT. 379 (1993) (noting incoherence of state action doctrine due to dismantling public/private distinction in post-1937 era combined with a contradictory continued commitment to notion of individual rights.)

115. GILMORE, *supra* note 1, at 111.

conferred by its competitor.[116] Of course, the reverse phenomenon is present as well. We observe in the law repeated efforts to generate empty spaces, often by means of replacing the indeterminacy generated by standards with the greater certainty generated by rules (such as safe-harbor rules).[117] But the very prominence of these efforts, and of the oscillation between standards and rules, illustrates the point of contraction as a pervasive phenomenon.

Contraction does not, of course, always proceed from a self-consciously "scientific" construction. Some contraction may be traced to conservative traditionalism of a self-consciously "moral" variety. Much can be traced to egalitarian commitments: the conferral of "power" by background entitlement tends strongly to render egalitarians hostile to empty spaces.[118] All, however, may be traced to an insistence upon articulate justification and a claim to authority in assessment of justification.

If this is so, how does it serve as a defense of empty spaces? It does so in two senses. First, as a descriptive matter, it undermines the realist claim that there are no empty spaces, for it makes no sense to deny the existence of the private while simultaneously substituting for some status quo an insistence upon justification and authoritative assessment. One does not substitute a proffered reality for a non-existent alternative reality. Second, it makes clear that the debate over empty spaces is normative. The anti-formalist has a normative agenda that cannot be defended in merely descriptive terms. So, too, of course, does the formalist, if commitment to empty spaces is accepted as a formalist precept.

IV. THE NORMATIVE DEBATE

What is the nature of this normative debate? The nature of the normative debate may be found in the following general criticism of formalism: by refusing to address consequences, formalism constitutes an abstract theology divorced from social need.[119] It seems to me that within this criticism are the roots of the fundamental disagreement. That disagreement entails two interrelated issues: competence and ambition.

---

116. *See, e.g.*, CAROL M. ROSE, PROPERTY AND PERSUASION, ESSAYS ON THE HISTORY, THEORY AND RHETORIC OF OWNERSHIP 199-225 (1994); Jason Scott Johnston, *Uncertainty, Chaos, and the Torts Process: An Economic Analysis of Legal Form*, 76 CORNELL L. REV. 341 (1991). An example of the latter phenomenon is the simultaneous presence of an individualistic disparate treatment theory and collectivist disparate impact theory in the law of Title VII. *See* PAUL N. COX, EMPLOYMENT DISCRIMINATION, ch. 6 (3d ed. 1999).

117. *See, e.g.*, S.E.C. Rule 506, 17 C.F.R. § 230.506 (amended 1989); REV. MODEL BUS. CORP. ACT §§ 8.60-8.63 (1984); UNIF. LIMITED PARTNERSHIP ACT § 303 (1976).

118. *See, e.g.*, SUNSTEIN, *supra* note 102, at 40-92 (status quo neutrality reflected in *Lochner* era non-neutral and unjust); Gerald E. Frug, *The Ideology of Bureaucracy in American Law*, 97 HARV. L. REV. 1276 (1984) (unjust power, for example, of corporate management).

119. POSNER, OVERCOMING LAW, *supra* note 5, at 398-99.

A. Competence

The notion that formalism is an abstract theology that refuses to address consequences obviously implies that there are better alternatives. It seems to me that formalism can be understood as a denial of this implication, and in particular a denial of the competence of legal actors either to resolve fundamental moral, political or social issues or to adequately predict and control social consequences. Its competitors, by contrast, affirm the capacity of law, of moral reasoning, or of scientific method to do just these things.

Recall that the formalist seeks his guidance from the concepts, rules, principles, etcetera he finds in the past practices of law, practices I earlier claimed nevertheless must inevitably have had some substantial relation to social practice even while not duplicating social practice. This source of legal decision is, by reference to the alternatives offered by anti-formalists, a quite modest one. It does not seek answers through the highfaluting techniques of analytical moral philosophy; it does not place its faith in the supposed expertise of administrative agencies; it does not suppose that social science is capable of achieving with the social what natural science has achieved with the natural. I submit that the claims to truth finding, prediction, control, and moral imperative one finds in these alternatives are far more extravagant than a simple claim to adherence to principles embedded in past practice. The alternatives display both high ambition—the ambition of improving society by reference to some philosophical, political, moral or economic precept—and a deep faith in the capacity of elites to employ rationality in service of this ambition.

Nevertheless, I do not wish to be understood as wholly rejecting criticism of formalist conceptualism. In particular, I do not believe that legal decision in hard cases can be thought of as compelled by past practice, even though that practice will substantially limit the alternatives. Indeed, I do not even believe that "reason" determines the choice between the alternatives thrown up by past practice. The skeptical realists and post-realists are, in my view, correct at least to this extent. The pretense of decision compelled by reference to principle may be a necessary pretense in such cases, but it is, I think, absurd to believe, as our legal culture asserts and purports to believe, that there are correct answers in hard cases, discoverable through reason.[120] This is particularly obvious when the hard case entails clashes between deeply felt political or moral commitments. There is simply no possibility of a rationally justified right answer in such cases.[121] This means, however, not only that right answers won't be found in legal principles. It also means they also won't be found in moral philosophy, economics or any other discipline or body of knowledge outside law.

I also do not wish to be understood as thinking consequences do not matter to what law is or should be; they obviously do matter. My points about the matter of consequences are that both formalists and anti-formalists exaggerate the degree to which formalist law ignores consequences in favor of principles and that ambitious consequentialist programs, like ambitious moral ones, should be

---

120. *See generally* PAUL F. CAMPOS, JURISMANIA: THE MADNESS OF AMERICAN LAW (1998).

121. *Id.*

greeted with a great deal of skepticism.

That formalist principle may be understood as utilitarian in character is suggested by the proposition that at least some common law doctrines were "efficient."[122] It is suggested by a Humean understanding of the "utility" of rules yielded by social practice to the extent these are incorporated in law.[123] It is suggested by a rule utilitarian, rather than act utilitarian version of proper consequentialist approach and by recognition that administrative cost, particularly the "cost" of irremediable official ignorance, is very high.[124] I do not here offer a utilitarian account of the common law, the form of law conceived by classical formalists, as *the* law. Others have done so.[125] I claim merely that formalist conceptualism and rule worship may have masked an underlying consequentialism, albeit one of limited ambition.

I greet more ambitious consequentialism with skepticism not because it lacks appeal. Economic analysis of law, a sophisticated form of consequentialism, seems to me the most intellectually appealing of extant alternatives. It is particularly attractive because it takes seriously, rather than merely paying lip-service, to the idea that there are two sides to every story: every benefit has a cost. Moreover, elements of that analysis have had the salutary effect of defeating naive consequentialism: the unfortunate belief that, by prohibiting some bad or requiring some good, the bad will be banished and the good will displace the status quo.[126] Nevertheless, we should also be skeptical of sophisticated consequentialism for the simple reason that we lack, and are likely to continue to lack, information necessary to it. Let me briefly explain this skepticism.

There are two distinct levels at which consequentialist prediction and weighing exercises might occur, although the distinction will be fuzzy in practice. One level may be labeled institutional. It entails assessment of the predicted costs and benefits of alternative institutional arrangements, particularly the alternatives of markets and governmental and non-governmental hierarchies.[127] The other may be labeled infra-institutional. It entails the adoption and use of the prediction of consequences and the weighing of costs and benefits as a method of decision within a given institution.[128]

---

122. *E.g.*, RICHARD A. POSNER, ECONOMIC ANALYSIS OF LAW 271-81 (5th ed. 1998).

123. DAVID HUME, ENQUIRY CONCERNING THE PRINCIPLES OF MORALS,§ III, pt. II (3d Selby-Biggs ed. 1975) [1777]; *see* HAYEK, LAW, LEGISLATION AND LIBERTY, *supra* note 4, at 113.

124. HAYEK, LAW, LEGISLATION AND LIBERTY, *supra* note 4, at 113.

125. *See* RICHARD ALLEN EPSTEIN, SIMPLE RULES FOR A COMPLEX WORLD (1995).

126. *E.g.*, Richard Craswell, *Passing On The Costs of Legal Rules: Efficiency and Distribution In Buyer-Seller Relationships*, 43 STAN. L. REV. 361 (1991).

127. *E.g.*, NEIL K. KOMESAR, IMPERFECT ALTERNATIVES: CHOOSING INSTITUTIONS IN LAW, ECONOMICS, AND PUBLIC POLICY (1994); OLIVER E. WILLIAMSON, THE ECONOMIC INSTITUTIONS OF CAPITALISM (1985); R.H. Coase, *The Nature of the Firm*, 4 ECONOMICA 386 (1937); R.H. Coase, *The Problem of Social Cost*, 3 J. L. & ECON. 1 (1960).

128. Thus, for example, judicial decision under "reasonableness" or "under all facts and circumstances" tests, where given a balancing of costs and benefits gloss, entails infra institutional

Consider, first, infra-institutional predicting and weighing. Hard formalist rules, at least those whose content creates or facilitates what I have called empty spaces, tend to allocate decision making authority to "private" or "market" institutions. If rationalist depictions of human behavior are correct, persons within these empty spaces then engage in prediction and weighing exercises. The hard rules that surround and support these empty spaces may often serve, or, at least, be explained as serving the function of compelling persons to consider, in their weighings, the goods and the bads inflicted by their actions on others. However, it remains the case that persons operating within such empty spaces have jurisdiction over prediction and weighing.[129] By contrast, anti-formalist "soft rules" or "standards" allocate this jurisdiction to governmental functionaries, to the extent that these personages have authority to make "all things considered" judgment. They will ultimately engage or threaten to engage in predicting and weighing. This is true, as well, however, of hard rules that direct particular outcomes and means of achieving those outcomes, for such rules deny or destroy empty spaces. The governmental functionaries who create such directive rules have engaged in an ex ante predicting and weighing in either naive or sophisticated versions. Prediction and weighing occurs, then, within distinct institutions and is therefore engaged in by distinct classes of persons.

Consider, now, prediction and weighing in the choice of institution. The

---

prediction and weighing. The economic interpretation of negligence is an obvious example. *E.g.*, WILLIAM M. LANDES & RICHARD A. POSNER, THE ECONOMIC STRUCTURE OF TORT LAW (1987). On the other hand, strict liability is not an alternative to prediction and weighing if this method is employed in identifying the party who will be strictly liable, as in analysis of the "least cost avoider." And negligence need not entail a regime of ongoing prediction and weighing if it is dominated in fact by rules. OLIVER WENDELL HOLMES, THE COMMON LAW 98-99 (M. DeWolfe Howe ed., 1963). *See* Stephen G. Gilles, *Negligence, Strict Liability, and the Cheapest Cost-Avoider*, 78 VA. L. REV. 1291 (1992); Stephen G. Gilles, *Rule-Based Negligence and the Regulation of Activity Levels*, 21 J. LEGAL STUD. 319 (1992).

129. A complication, however, is the matter of remedy. In the standard analysis, "property rule" remedies (such as injunctions and, perhaps, specific performance orders) force questions of allocation into market or contracting institutions and, therefore, would be favored in the "formalist" scheme I am depicting. Guido Calabresi & A. Douglas Melamed, *Property Rules, Liability Rules, and Inalienability: One View of the Cathedral*, 85 HARV. L. REV. 1089 (1972). (This would also be true of contractions of liability, the expansion of the realm of *damnum absque injuria*, because a dismissal order is the partial analogue, for the complaining party, to an injunction against the responding party). Also in the standard analysis, liability rule remedies (damages) are employed where contracting is obviated by transaction costs, and damages are prices. The difficulties with damages are that they "substitute" governmentally determined objective estimates of cost for a fundamentally subjective experience of cost, rendering them prone to error and unpredictable. Governmental pricing of behavior may be said to leave choice jurisdiction in the hands of "private actors," as, for example, in the notion that "efficient breach" justifies expectation damages. But it also is governmental pricing, so there can be no assurance that the prices set reflect those that would be subjectively demanded. Perhaps more importantly, I submit that these prices are not predictable ex ante, so the incentive function justifying these prices is in doubt.

allocation of jurisdiction might be decided on the basis of predicting and weighing. One might say, for example, that transaction costs in a particular context preclude appropriate private decision within an empty space and that the distortions of interest group politics are unlikely to be present in this context, so, on balance, jurisdiction should be allocated to a judicial, "political," "public," or "administrative" institution. Alternatively, one might predict that transaction costs in a particular context are low and governmental information costs high, so, on balance, jurisdiction to engage in infra-institutional predicting and weighing should be allocated to the empty space.

With these preliminaries out of the way, let me return to the matter of skepticism about competence as a justification for formalism, addressing, first, sophisticated prediction and weighing as a means of doing law and, second, such prediction and weighing as a basis for allocating decision-making jurisdiction.

By "sophisticated prediction and weighing as a means of doing law," I mean the use of these methods by legal authorities in making particular decisions, and, therefore, assume allocation of choice making jurisdiction to governmental authority. I also again mean, however, the use of these methods in formulating hard rules of a command and control variety: rules, formalist in their hard form, but anti-formalist in their rejection of empty spaces. A rule that directs ends and means is functionally equivalent to an "all things considered" decision by a governmental functionary, for, in both instances, it is a governmental institution that determines particulars. The phenomena differ only in time (ex ante or ex post) of governmental decision.

The reasons for skepticism are many and have been repeatedly offered by others. Let me, however, briefly rehearse some of these reasons: (1) The, ironically, formalist method of prediction employed by sophisticated prediction and weighing, which is rigorous deduction from the rationality and scarcity postulates, misspecifies the complex character of human behavior.[130] (2) The specification of particular motivations as the ends sought through means-ends rationality too often misspecifies the complexity of human motivation.[131] (3) The objective prices necessarily postulated in weighing exercises either ignore or are poor proxies for the reality of the subjectivity of cost.[132] (4) The commitments of the analyst therefore necessarily color objective price estimates.[133] (5)

---

130. *E.g.*, HOWARD MARGOLIS, SELFISHNESS, ALTRUISM AND RATIONALITY (1982); RICHARD H. THALER, THE WINNER'S CURSE: PARADOXES AND ANOMALIES OF ECONOMIC LIFE (1992); Christine Jolls et al., *A Behavioral Approach to Law and Economics*, 50 STAN. L. REV. 1471 (1998).

131. *E.g.*, Amartya Sen, *Rational Fools: A Critique of the Behavioral Foundations of Economic Theory*, 6 PHIL. & PUB. AFF. 314 (1977).

132. *E.g.*, JAMES M. BUCHANAN, COST AND CHOICE (1969); Friedrich A. Von Hayek, *Economics And Knowledge*, *in* FRIEDRICH A. VON HAYEK, INDIVIDUALISM AND ECONOMIC ORDER 33 (1948). For interesting arguments regarding the implications of subjectivism, see Gregory Keating, *Reasonableness and Rationality in Negligence Theory*, 48 STAN. L. REV. 311, 337-41, 367-73 (1996); Gary Lawson, *Efficiency and Individualism*, 42 DUKE L.J. 53 (1992).

133. Duncan Kennedy, *Distributive and Paternalist Motives in Contract and Tort Law, with*

Empirical evaluation of the hypotheses generated by the exercise most often does not occur.[134] (6) When empirical testing does occur, the tests employed are insufficiently sensitive; so, while they may produce results consistent with a tendency with which the hypothesis is also consistent, they cannot satisfy a falsifiability criterion.[135] (7) When empirical testing occurs and generates suggestive results, it is always subject to methodological and interpretive challenge, and, most often, these challenges are sufficiently weighty to preclude reliance. Therefore, there is typically an unsurprising positive correlation between prior political or moral commitment and interpretation of empirical findings.[136] (8) Finally, the analytical apparatus is so "rich," or perhaps porous, that it permits competing and inconsistent plausible hypotheses about behavior,[137] again often correlated with prior commitment, and choice between these

---

Special Reference to Compulsory Terms and Unequal Bargaining Power, 41 MD. L. REV. 563, 597-604 (1982); Duncan Kennedy, *Freedom and Constraint in Adjudication: A Critical Phenomenology*, 36 J. LEGAL EDUC. 518 (1986).

134. The best evidence of this phenomenon are the pleas of advocates for more empirical research. *E.g.*, POSNER, *supra* note 22, at 164, 217.

135. For example, empirical evidence supports the proposition that "incentives matter." *E.g.*, POSNER, *supra* note 122, at 220-24 (providing evidence indicating that tort liability reduces accidents). The more difficult issue, however, is whether a particular form of incentive matters, and, more specifically, whether attempts at precision in formulating legal incentives matter. This may be doubted. *See* Gary T. Schwartz, *Reality in the Economic Analysis of Tort Law: Does Tort Law Really Deter?*, 42 UCLA L. REV. 377 (1994).

136. There are, of course, numerous examples; I offer the following as representative: *Compare* William G. Bowen & Derek Bok, *The Shape of the River: Long Term Consequences of Considering Race in College and University Admissions* (1998), *with* Stephen Thernstrom & Abigail Thernstrom, *Reflections on the Shape of the River*, 46 UCLA L. REV. 1583 (1999). *Compare* Terrrance Sandalow, *Minority Preferences Reconsidered*, 97 MICH. L. REV. 1874 (1999) and Terrance Sandalow, *Rejoinder*, 97 MICH. L. REV. 1923 (1999), *with* William G. Bowen & Derek Bok, *Response to Review by Terrance Sandalow*, 97 MICH. L. REV. 1917 (1999). Additionally, *compare* Roberta Romano, *The State Competition Debate in Corporate Law*, 8 CARDOZO L. REV. 709 (1987), *with* Elliott J. Weiss & Laurence J. White, *Of Econometrics and Indeterminacy: A Study in Investors' Reactions to "Changes" in Corporate Law*, 75 CAL. L. REV. 551 (1987); *cf.* Adrian Vermeule, *Interpretation, Empiricism, and the Closure Problem*, 66 U. CHI. L. REV. 698 (1999) (maintaining empirical inquiry often unable to answer questions it addresses at reasonable cost and within useful period of time).

137. POSNER, *supra* note 2, at 363-67. *See* Henry N. Butler, *The Contractual Theory of the Corporation*, 11 GEO. MASON L. REV. 99 (1989) (criticizing "misapplications" of theory). For example, consider the matter of insider trading prohibition and the many ingenuous efforts at justifying it in economic terms in face of the standard economic critiques of the prohibition. For an overview of this debate from a critical viewpoint, see STEPHEN M. BAINBRIDGE, SECURITIES LAW: INSIDER TRADING 125-73 (1999). For an example of ingenuous effort, see the work of my colleague, Nicholas Georgakopoulos. *Insider Trading as a Transactional Cost: A Market Microstructure Justification and Optimization of Insider Trading Regulation*, 26 CONN. L. REV. 1 (1993).

hypotheses cannot be made within the spirit of "scientific" inquiry absent more powerful empirical mechanisms than we possess or are likely in the future to possess.

What of prediction and weighing as a method of allocating decision-making authority? The issue here is who should decide, in particular, which institution should decide. It may seem that I have already loaded the argument in favor of "private" realms or market institutions by expressing skepticism about the prediction and weighing capacities of governmental actors, but this is not *yet* quite the case. If governmental actors are poor predictors and weighers, so, too, may be private actors. So the question of institutional allocation is distinct from the question of method assuming an allocation. The question of prediction and weighing as a method of determining an appropriate allocation is, likewise, distinct from the question of this method employed as a device for reaching particular decisions.

The issue with respect to allocation is, presumably, that of relative institutional competence: which institution is most likely to make the best decisions? Unfortunately, however, this question assumes an answer to a further underlying question: what is meant by "best"? A prediction and weighing method of answering the allocation question would seem to assume a welfarist criterion as an answer to this underlying question, quite possibly an efficiency criterion. On this assumption, the allocation question becomes: which institution is most likely to generate "efficient" outcomes?[138]

Persons who approach legal issues from the perspective of this allocation question tend to do so by identifying various defects in the institutions in question, usually defects that serve as obstacles to efficiency.[139] Markets or private contracting institutions are afflicted with "transaction costs." Political institutions and administrative agencies are affected with the rent-seeking evils of interest group politics. Courts and juries are afflicted with an inability to initiate action, costly processes, and substantial questions of competence. The method of prediction and weighing in assessing the allocation question is therefore one of predicting relative institutional performance and weighing the force of these defects in particular contexts.

The method, when applied to the question of allocation, potentially suffers from the problems recounted above when applied to actual decisions given an allocation. In particular, it would suffer from these problems if it purported to identify with precision the monetary or other values to be assigned the costs and benefits of alternative institutions. This, however, is rare. The more typical exercise in this form of analysis is unquantified description. The analysis therefore tends to rely upon what I term "knowable tendencies" or

---

138. This again, however, is not the only possible criterion. One might seek to make predictions about which institution is best able to effect egalitarian outcomes. KOMESAR, *supra* note 127, at 34-49.

139. *Id.* at 53-152; Daniel H. Cole, *The Importance of Being Comparative*, 33 IND. L. REV. 921 (2000).

generalizations about human behavior and not upon unknowable particulars.[140] Moreover, analysis of comparative institutional competence is Hayekian in spirit, for it recognizes that institutional capacity is the central question.

Nevertheless, there are reasons to be skeptical of the method of prediction and weighing when applied to the question of allocation of jurisdiction, even when the method relies upon general tendencies and eschews quantification of particulars. One reason is that the historical, perhaps even systematic, tendency has been one of identifying defects in one institution while assuming that its alternatives are free of defects.[141] This is a problem that may be overcome in theory; good comparative analysis can be substituted for bad comparative analysis.[142] The tendency to bad analysis is nevertheless a tip-off to a second problem. In the absence of an adequate mechanism for quantifying cost and benefits, a mechanism I have been suggesting is not in the cards, prediction and weighing will reflect prior commitments to a degree that the exercise will merely confirm these priors. If my prediction is incorrect, if there are at least some cases in which unquantified reliance upon general tendencies can yield predictions free of the taint of prior commitment, there is a third problem. We will most often discover both that the defects of alternative institutions are highly correlated and that their values, while unquantified, are probably high. The result is that we are left, or, most often will be left, with no clear answer to the question of relative institutional competence.[143] In the absence of an objective answer, we will again fall back on our priors, appearing now as presumptions left unrebutted by the exercise.

My final reason for skepticism is that exercises of this sort purport to proceed from outside the institutions examined, as if the analyst, from this outside stance, were in a position to allocate jurisdiction free from the defects she detects in these institutions. This, of course, is pure fiction. There is no single, conscious, impartial, and adequately knowledgeable entity standing outside the subject matter and possessing authority to allocate. The fiction is useful as thought experiment. But it is pernicious if we lose track of the fact that the choosers of institutions are our existing highly imperfect institutions – the institutions subject to the failures neoinstitutionalists identify.

B. Ambition

Although I have mentioned the matter of ambition, I have not yet directly addressed it. I said above that ambition is one of the two interrelated sources of normative disagreement about formalism. I derive this from the claim that formalism fails to respond to "social need." The implied ambition is that of satisfying or resolving social need. Just what might be meant by "social need"?

---

140. These, at least, are my impressions. *Cf.* POSNER, OVERCOMING LAW, *supra* note 5, at 426-37 (describing neoinstitutional theory's rejection of economic formalism).

141. Coase, *supra* note 127.

142. Cole, *supra* note 139.

143. James E. Krier & Stewart J. Schwab, *Property Rules and Liability Rules: The Cathedral in Another Light*, 70 N.Y.U. L. REV. 440 (1995).

There are distinct conceptions of the function of law and of the "social need" functionally served by law.

Classical formalists conceived of law as the common law.[144] Important features of the common law, as it was addressed by the classical formalists, were that it was decentralized, transactional, corrective, historical, derivative, status neutral, and in an important sense purposeless.[145] By "decentralized," I mean that the common law is the product of a series of decisions in concrete cases by distinct judges. It has no identifiable, central author, and therefore resists both positivism's demand for a sovereign source and legal realism's positivist fixation on the judge as a declarer, rather than a follower, of law.[146] By "transactional," I mean that its focus and subject matter is upon particular transactions, whether voluntary or involuntary, between individuals. By "corrective," I mean that it is concerned about the making, or not, of "wrong moves" by individuals within such transactions. Indeed, it assumes and preserves a status quo by addressing wrong moves that have disturbed the status quo. By "historical," I mean that it addresses past transactions. While it thereby establishes guidance (or rules) for future transactions, it does not in a broad legislative sense purport to prospectively legislate the future in service of a defined collective objective. By "derivative," I mean that it is derived from social practice or common morality, in the way indicated by my earlier discussion of intuitionism.[147] It is not, then, directive of social practice in the way that a command originating from a source alien to social practice is directive. By "status neutral," I mean that it is individualistic in the sense that the actions of individuals, not their status or group membership, count. It is therefore "general," in the sense that it is *formally* neutral. By "purposeless," I mean that it does not, at least directly, seek to achieve some consciously articulated collective objective or end-state.

If this is correct as a depiction of the common law, classically conceived, it is decidedly non-functional when function is understood as service to consciously articulated social end-states, and it decidedly fails to serve social need when this need is defined in terms of such end-states. But this does not preclude it from being functional in the sense of enabling persons to identify with whom and by what means they may transact with others in service of their individual

---

144. *E.g.*, Grey, *supra* note 10, at 34-35.

145. I rely, in what follows in the text, upon: Barry, *supra* note 33; Gjerdingen, *supra* note 33, at 876-83; and Mashaw, *supra* note 33, at 1153-59.

146. *See* JOHN CHIPMAN GRAY, THE NATURE AND SOURCES OF THE LAW 82, 91 (1916). *But see id.* at 116. I deem that strand of legal realism that emphasizes the judge as a source of law "positivist" in that positivists are supposed to be committed to a sovereign source of law. Realists could, of course, either favor the judge as a sovereign (Llewellyn) or disfavor that source (as in those realists who preferred rule by expert administrative agencies). William W. Bratton, *Berle and Means Reconsidered at the Century's Turn*, 26 J. CORP. L. 737, 741-50 (2001).

147. *See supra* notes 35-42 and accompanying text. *Cf.* Donald H. Gjerdingen, *The Coase Theorem and the Psychology of Common-Law Thought*, 56 S. CAL. L. REV. 711 (1983) (indicating classical common law thought appeals to normative intuitions of lay persons).

preferences.[148]

The obvious objection to equating formalism with this depiction of the classical common law is that classical formalism's alleged "top-down" autonomous conceptualism—its commitment to deriving legal answers from legal principal—appears inconsistent with a decentralized, "bottom-up" common law, a common law built up from resolution of particular actual cases.[149] The "scientific" aspirations of classical formalism – its attempt to select the one correct rule from what Langdell thought was the "useless" jumble of the common law[150]—may be viewed as one well within a centralized, directive, and prospectively legislative tradition incompatible with this depiction of the features of classical common law.[151] Indeed, Grant Gilmore's conception and critique of formalism may perhaps best be read as hostility to this ambitious, directive depiction. Gilmore's apparent understanding of his anti-formalism was one of favoring fact sensitive, almost ad hoc judgment, or, at least, judgment tied only loosely to principle, and one, following Llewellyn, relying heavily on social practice.[152] Nevertheless, I think a formalist label warranted. Let me supply four reasons for this view.

First, it is important to again recognize that the classical formalists were engaged in an inductive project of identifying principles that would reconcile, systemize, and render coherent the common law. The source of their principles was common law precedent.[153] To systemize and rationalize is to centralize in a sense, but, to the extent that the formalist project rested upon the products of a decentralized process, and sought to be true to these products,[154] it remained decentralized in its origin. In short, the classical formalists sought to restate, in coherent form, the traditions of the common law. Now it is true that they are also typically understood as rigidifying the common law, as exaggerating its

---

148. That is, the law of property, contract and tort may be understood as concerned with enabling exclusion of others (private property), enforcement of promised exchanges (contract) and establishing a knowable line between permissible and impermissible externalization (tort), all on the assumption of a classically liberal (or, if one wishes, "atomized") order. *See* HAYEK, LAW, LEGISLATION AND LIBERTY, *supra* note 4, at 112-15.

149. POSNER, OVERCOMING LAW, *supra* note 5, at 172-73. *See* MELVIN ARON EISENBERG, THE NATURE OF THE COMMON LAW 146-61 (1988); ANTONIN SCALIA, A MATTER OF INTERPRETATION 3-14 (1997); SCHAUER, *supra* note 49, at 174-81.

150. CHRISTOPHER COLUMBUS LANGDELL, A SELECTION OF CASES ON THE LAW OF CONTRACTS v-vii (1871), *quoted in* STEPHEN B. PRESSER & JAMIL S. ZAINALDIN, LAW AND JURISPRUDENCE IN AMERICAN HISTORY 734-36 (4th ed. 2000). *See* Grey, *supra* note 10, at 11 n.35, 24-27.

151. Indeed, Hayek at one point so viewed it. HAYEK, LAW, LEGISLATION AND LIBERTY, *supra* note 4, at 106.

152. GILMORE, *supra* note 1, at 108-11.

153. Grey, *supra* note 10, at 24-32.

154. This, in the case of the classical formalists was a condition arguably not met. A standard objection to their efforts was their selective treatment of caselaw and failure, therefore, to recognize what was "really" going on. *E.g.*, Walter Wheeler Cook, *Williston on Contracts*, 33 ILL. L. REV. OF NW. U. 497 (1939) (reviewing the Williston treatise).

coherence, as falsely supposing its completeness, and as misidentifying the mechanism of decision as deduction from principle rather than "utility," "situation sense," or "felt need."[155] If it is true, however, that utility, situation sense and felt need were the true mechanisms that brought about the precedents from which the classical formalists derived their principles, it is difficult to believe that these principles were independent of the mechanisms. They more plausibly reflected the mechanisms.

Second, the noted features of common law are, rather precisely, the opposites of the features of law advocated by many critics of classical formalism—legal realists, post-realists, and pragmatic instrumentalists. For many of the critics, proper law is centralized, patterned, distributive, forward looking, directive, status conscious, and purposive.[156] It is "centralized" in that realists were obsessed with the judge as an author or maker of law (as opposed to applier or interpreter of law) and, at least in post-realist practice, favored legislative direction and the supposed expertise of administrative agencies, particularly at the federal level. It is "patterned," "distributive," and "forward looking" in that it is viewed as an instrument for conforming classes of conduct to articulated collective objectives and therefore for reform of the status quo. It is "directive" in that law is an instrument for reforming social practice on the basis of principles or policies derived independently of that practice. It is "status conscious" in that it focuses upon groups and deems these important. It is therefore not general in that the legal rights and obligations it recognizes are dependent upon status or context. It is "purposive" in that realist and post-realist law is an instrument for achieving collectively articulated "social" ends.

These features of realist aspiration have, of course, at least partially become features of current law—the law of the "administrative state."[157] This is true not merely in the law as interpreted and enforced by administrative agencies, but also within the common law itself. The law of torts, of contract, of property are now largely conceptualized in these instrumental terms both within academia and

---

155. *See* HOLMES, *supra* note 128 (felt necessities of the time); HUME, *supra* note 123 (utility); LLEWELLYN, *supra* note 19, at 268 (situation sense);.

156. I here again rely upon Gjerdingen, Mashaw, and Barry, *supra* note 33.

157. *See* G. EDWARD WHITE, PATTERNS OF AMERICAN LEGAL THOUGHT 99 (1978) (realism as intellectual analog to the New Deal). Professor Duxbury argues that the New Deal (and, by implication, post-New Deal administrative state) were not reflections of legal realist jurisprudence on the ground that the legal realists, as academics, failed to develope a theory of administrative law. DUXBURY, *supra* note 7, at 153-58 (nevertheless citing Roscoe Pound and Jerome Frank for the proposition that legal realism and the New Deal were linked). While it is true that the legal realists, as academics, focused on "private law," and so offered a perspective on the common law opposed to the classical characterization, it is precisely, I submit, the realist perspective that was later reflected in New Deal and post-New Deal regulatory programs. *See id.* at 7, 78 (realism in part a response to laissez faire); *id.* at 79-82 (realism as resort to social sciences with object of social control); *id.* at 97-111 (realism as reflecting institutional economics, particularly its egalitarian themes).

within the profession.[158] Similarly, contemporary depictions of the common law, in contrast to the rigid traditionalism of classically formalist depictions, tend to treat rules as mere guideposts to decision by a governmental functionary in instrumental service of socially desirable ends.[159] In short, formalism's antagonist was and remains a set of beliefs at the core of which is the conviction that human societies can and should be consciously planned or constructed. It is in this set of beliefs that another, more ambitious understanding of function and of social need are evident and to which formalism is "blind" or antagonistic.

Third, classical formalism's scientific pretensions were, as Professor Grey has demonstrated, quite unlike the scientism of pragmatic instrumentalism.[160] Science, for classical formalists, entailed the paradigm of a closed logical system. The objective was to render law on the model of geometry. The scientism of formalism's antagonist is closer to more current understandings of science, with its emphasis upon hypothesis and empirical verification, fondness for experimentation, and objective of human control over natural phenomena. Langdell's science of law was a science of conceptual consistency. Realism's science of law was a science of conscious, purposeful social control. There is, then, a distinct lack of ambition in formalist science, at least when compared to its competitor.

Finally, it is not necessary to a contemporary formalism that even classical formalism's ambitions be duplicated. Given my concessions that law as geometry is implausible and that right answers in hard cases cannot be uncontroversially resolved through reason,[161] classical formalism's pretensions to science should be abandoned. What might then remain, however, could very much be in the spirit of the classical common law. For example, dominant contemporary views of the common law as a fluid process might give way to more rigid views, views in which stare decisis would be taken more seriously, attempts at distinguishing precedent would be looked upon with more skepticism, and arguments from social or economic change would be viewed with suspicion.

It is this comparative lack of ambition I wish to equate with formalism as a more contemporary project and with a contemporary formalist rejection of "social need" more ambitiously defined. It should be apparent that comparative lack of ambition is related to skepticism about methodological capacity. I think skepticism about ambitious method leads to skepticism about, indeed antagonism toward, the idea of a collectively specified social end-state as objective, and law as means to this objective. The reasons may be found in the tradition of Burkean conservatism, summarized in the law of unintended, but unquantifiable consequences and partially justified by our recent historical experience with the grotesque evils, grounded in ambition, that enjoyed too often and for too long the

---

158. Thomas C. Grey, *Hear The Other Side: Wallace Stevens and Pragmatist Legal Theory*, 63 S. CAL. L. REV. 1569, 1590 (1990); Summers, *supra* note 8.

159. *E.g.*, EISENBERG, *supra* note 149; SCALIA, *supra* note 149, at 3-14; SCHAUER, *supra* note 49, at 174-81.

160. Grey, *supra* note 10 at 16-20.

161. *See supra* notes 20-21, 120-21 and accompanying text.

support of an intelligentsia confident of its capacities.[162]

C. Formalism and Politics

Let me conclude my account of the debate between formalists and anti-formalists by addressing an obvious question: Is formalism a political program? I have been defending formalism as contract dominated, common law permeated, with empty spaces. Is my version of formalism simply, then, a species of libertarian or classically liberal political commitment?

It is surely the case that critics of formalism have depicted it as substantive, as a species of conservative or reactionary ideology.[163] *Lochner v. New York*,[164] in keeping with this depiction, is, for example, often deemed an example of formalism. It seems also reasonably clear that American legal formalism is historically associated with free market, laizze faire or libertarian positions.[165] On the other hand, *Lochner* is not in fact an example of a formalist mode of adjudication; it is an example of the use of a balancing test, albeit one employed in service of a laissez faire agenda.[166] Perhaps formalist methods, like anti-

---

162. I am not equating legal realism or pragmatic instrumentalism with National Socialism or Communism. Nor am I suggesting that realism or pragmatism inevitably result in such evils. I am, however, suggesting that excessive ambition in law can be dangerous. *Cf.* POSNER, OVERCOMING LAW, *supra* note 5, at 153-59 (recognizing, on the basis of INGO MÜLLER, HITLER'S JUSTICE: THE COURTS OF THE THIRD REICH (Deborah Schneider trans., 1990), that it was not legal positivism, but a rejection of positivism, that explains the behavior of German judges in the Nazi era); Cass R. Sunstein, *Must Formalism Be Defended Empirically?*, 66 U. CHI. L. REV. 636, 636-37 (1999) (same).

163. HOROWITZ, *supra* note 2; *cf.* SUNSTEIN, *supra* note 102, at 46-92 (critique of status quo neutrality); *but see* SUNSTEIN, *supra* note 5, at 118-20 (rejecting link between rule of law and free markets).

164. 198 U.S. 45 (1905).

165. DUXBURY, *supra* note 7, at 25-32; HOROWITZ, *supra* note 2, at 33-39, 142, 193, 200.

166. POSNER, OVERCOMING LAW, *supra* note 5, at 284; *cf.* HERBERT HOVENKAMP, ENTERPRISE AND AMERICAN LAW 1836-1937, at 172-75 (1991) (generally rejecting formalism as explanation of substantive due process). Perhaps the best argument for deeming *Lochner* a formalist decision is the claim that the constitutional concept of "liberty" does not compel freedom of contract, so the justices in *Lochner* were "dishonest" in not justifying their claim that this freedom was constitutionally protected. *See* SUNSTEIN, *supra* note 102, at 45-67; Schauer, *supra* note 6, at 514.

There are a number of difficulties with this contention. First, it does not explain why the non-economic "freedoms" recognized by post-New Deal constitutional law as derivable from "liberty" or other constitutional generalizations are not subject to the same claim. Granting that much ink has been spilled in attempted justification, no uncontroversial, ironclad argument supports these freedoms. Second, whether any given freedom is necessarily entailed by "liberty" depends upon whether the community believes it is so entailed. In a heterogenous community, consensus is unlikely. This implies that (1) *Lochner* did not unjustifiably derive contractual freedom from constitutional liberty given the beliefs of a community; it merely failed to recognize heterogeneity of belief and (2) this justification and failure support and infect currently recognized constitutional

formalist methods, may be employed to serve multiple political masters.

It seems to me, in fact, both that the various interpretations I have given formalism can operate independently of each other and that at least the law as rules and law as conceptualism interpretations of formalism can be independent of substantive political commitment. It is quite possible to formulate rigid rules on quite instrumentalist grounds and it is quite possible to deem rigid rules the most pragmatic means of achieving "social objectives." It seems to me, moreover, that much "left-wing" or "progressive" legal analysis warrants a formalism as conceptualism label. Substituting egalitarian conceptions of equality for libertarian conceptions of liberty is not an escape from conceptualism.[167] A good portion of consequentialist analysis is employed as "right-wing" or "conservative" rebuttal of "left wing" or "progressive" conceptualism.[168] The association of formalism with the right and anti-formalism with the left may therefore rest on historical contingency. So formalism and anti-formalism may simply be tools or weapons of convenience, with no necessary connection to any substantive political commitment.

Nevertheless, there is a case for thinking those critics of formalism who associate it with conservative or libertarian political commitments are largely correct. It is a case of affinity, and, perhaps, a case for the proposition that formalist form may be a necessary, though not sufficient, condition for implementing these commitments.

---

freedom. Therefore, (3), either the claim of dishonesty must fail or it must be applied to all controversial constitutional adjudication.

The claim that *Lochner* was "dishonest" is not, in my view, aided by the claim that it relied upon a "legally constructed" baseline as (falsely) neutral. This is my view for two reasons. First, it does not follow from the contention that the court relied upon a common law baseline (or that it sought to elevate the common law to constitutional status) that this baseline was consciously planned. It therefore does not follow that conscious planning of a new baseline, even given that some baseline is required, is justified. The common law and conscious, purposive planning entail distinct processes with distinct assumptions about human capacity. Second, if the alternative to a common law baseline is "deliberative democracy," it should be apparent by now that "deliberative democracy," as practiced, is perverse, or, at least, that it would not be unreasonable for a contemporary community to believe that it is perverse, given what we know from the "public choice" literature and given what we know of the electorate's ignorance. *Compare* Daniel A. Farber & Philip P. Frickey, *The Jurisprudence of Public Choice*, 65 TEX. L. REV. 873 (1987) (moderate criticism of public choice theory) *with* Michael DeBow & Dwight Lee, *Understanding ane Misunderstanding Public Choice: A Response to Farber & Frickey*, 66 TEX. L. REV. 993 (1988) (defense of public choice theory). *See, e.g.*, Samuel DeCanio, *Beyond Marxist State Theory: State Autonomy in Democratic Societies*, 14 CRITICAL REV. 215 (2002); Reihan Salam, *The Confounding State: Public Ignorance and the Politics of Identity*, 14 CRITICAL REV. 299 (2002).

Of course these musings suggest that *Lochner* was a formalist decision in precisely the sense that it relied upon a common law baseline and, if my earlier contentions are correct, that this baseline is a fundamental assumption of formalism.

167. POSNER, OVERCOMING LAW, *supra* note 5, at 271-86.

168. *E.g.*, POSNER, *supra* note 122, at 361-75, 514-18.

If we begin with skepticism about conscious, purposive governmental direction, it should be apparent that the various features of formalism I have postulated "fit" that skepticism at least in the sense that they are partial strategies for implementing it. The autonomy of law, in the form of traditionalist conceptualism, protects law from the ambitions of science (as science is *now* understood), and, therefore, society from law as constructivist social science. This autonomy serves also to protect law and society from the threat posed by anti- formalist, pseudo-scientific ideologies, ideologies illustrated by the decidedly anti-formalist examples of National Socialism and fascism in the last century.[169] This protection assumes that the concepts employed are "liberal," in the old, non-socialist, sense of the term, so the protection afforded may be historically contingent, but conceptualism, once this contingency is met, is a vehicle for avoiding a managed society.

Rigid rules provide determinate guidance, enabling coordination. If employed for purposes of coordinating individual behavior assumed to have been undertaken pursuant to diverse private ends, such rules enable empty spaces. This "if" is another contingency, for rigid rules may be employed to frustrate or preclude such a pursuit and to direct behavior in service of collectively formulated public ends. The Code of Federal Regulations is, after all, full of rigid-looking rules. Again, however, if this contingency is met, a rigid rule preference is a means by which the empty space becomes viable.

Perhaps, however, I have mischaracterized the political sides in this story. Consider the possibility that the debate is between authoritarians and anti-authoritarians. Given this way of looking at matters, my contention that skepticism about law justifies formalism will seem particularly ironic. On more standard accounts, formalism is grounded upon and expresses authoritarian certainty. This, recall, was Gilmore's perception: Formalism's conceptualistic abstractions, grounded in the dead hand of the past, ignore the particularized realities, the situation-specific needs and expectations of real people.[170] Classical formalists like Langdell ignored the operative facts of real cases in favor of their preferred principles, so formalism resembles the centralized directives of a distant commissar. One might respond that it is the administrative state, the culmination of legal realist thought, that better fits this commissar charge, but this rejoinder won't work against Gilmore; he had, or said he had, no sympathy for the administrative state and claimed that formalists and legal realists had in common both scientism and a lamentable belief in implementable truth.[171]

This brings me to the original question posed in this essay. I, largely following Hayek, have depicted formalism, or at least a version of formalism, as a strategy for minimizing law for anti-authoritarian reasons. Gilmore attacks formalism on the basis that it is an authoritarian conception of law.[172] How might

---

169. *See supra* note 162; *see also* Guido Calabresi, *Two Functions of Formalism*, 67 U. CHI. L. REV. 479 (2000).

170. GILMORE, *supra* note 1, at 41-56.

171. *Id.* at 100-01.

172. *Id.*

this conflict be explained? One clear possibility is that one of us is wrong in our understanding of formalism, or, perhaps more plausibly, that we have distinct interpretations of an amorphous concept. Another possibility is that this conflict reflects a deeper and more fundamental conflict between conceptions of what it means to be anti-authoritarian.

I think this second possibility is, in fact, a probability. There is a deep, fundamental conflict in perception. But I do not here attempt to diagnose its origins. Instead, I will attempt to point out some of its manifestations. One such manifestation is the distinction between an ex ante and ex post conception of law.[173] Formalism, as I have depicted it, is very much within the ex ante conception. Its anti-authoritarian strategy is that of providing a set of knowable rules in service of empty spaces human interaction.[174] "Freedom" falls out of the ability to know what to do to achieve one's ends through compliance with these knowable rules. Rules are therefore ex ante guides to behavior. Gilmore's dispute-centered version of law is, by contrast, one within the ex post conception. As I read him, he was concerned about what to do after the fact, and he answered with a version of all things considered, contextualized judgment. I take it that he wished to tie this judgment, through fact sensitivity, or "situation sense" to some version of cultural expectation. If so, it would not be rules or even common law precedents, but the capture of contextualized expectations that would generate, almost as an afterthought, any ex ante predictability.

Consider, in particular, Gilmore's anti-formalist rhetoric—the claim that formalism's abstractions impose themselves on real world, situation specific needs and expectations. This view makes perfect sense to anyone who places himself in the position of the judge, for example, in the imaginings of the legal academic. It makes sense because anyone with decent instincts will want a resolution of a dispute that seems to him just, all things considered. Hard formalist rule worship will therefore seem indecent. But this is the view of authority, of the person who has or wishes to have responsibility for decision. The point of "indecent" formalism is that it allocates jurisdiction for decision elsewhere.

Gilmore might respond by citing rule skepticism. If it is true that rules cannot themselves constrain, if all things considered judgment is inevitable and

---

173. *See* Robert E. Scott, *Chaos Theory and the Justice Paradox*, 35 WM. & MARY L. REV. 329 (1993) (discussing the justice paradox as tension between doing justice in particular case and regulation of future).

174. Can this assertion be reconciled with my transactional/historical depiction of classical common law, *supra* text and notes 144-48. It can, on the following grounds: For the law to be historical and transactional does not mean that it must be concerned with justice between the parties to a particular past transactional event on an all things considered basis. In the formalist version of historical and transactional justice, it means instead that law is concerned with identifying wrong moves as these are defined by knowable rules. Similarly, an ex ante perspective, one that seeks to establish guidance for the future, need not entail an effort to plan means of achieving a collectively determined end-state. In the formalist depiction, ex ante means simply the establishing of knowable rules for engaging in future transactions between individuals.

merely pushed underground by a norm of justification by reference to rule, formalist hopes are obviously at risk. And if the real constraint is attitudinal—the formalist judge's good faith effort to be a formalist and Gilmore's judge's good faith effort to be a wise interpreter of cultural expectation—the formalist cannot viably claim he has a better means of constraining ambition.

Perhaps this is correct, but I do not believe that it is to a degree that would obviate the claim that formalism's constraints on ambitious law are superior to Gilmore's reliance on official wisdom. If I am correct in believing extreme rule skepticism unjustified, formalism's constraints provide a basis for disciplining decision and a benchmark for critique. An appeal to open-ended wisdom does not.

CONCLUSION: IS FORMALISM LIKELY?

I have thus far argued that formalism is both viable and, at least to me and perhaps a few others, attractive. I will close by addressing the question whether it is likely—whether, that is, there is a reasonable prospect that it will triumph.[175] My answer is no. I do not mean by this answer either that formalism is wholly absent from American law or that it will disappear from American law. It is both present and enjoying in some contexts a resurgence. Nevertheless, I think the prospects for its triumph unlikely for two sets of reasons.

First, underlying formalism are a set of values, or, perhaps, personality traits, that are largely absent in contemporary America, particularly within the intelligentsia. Formalism requires restraint in the form of a tolerance of apparent injustice, apparent absurdity, even apparent evil. I say "apparent" because injustice, absurdity and evil are more often than not controversial characterizations rather than reflections of consensus, because the benefits of correcting these bads, even where there is consensus that they are bads, are always accompanied by costs to legitimate interests and values, because these costs are often ignored and often thoughtlessly denigrated, and because the terms

---

175. A fair question is what would such a triumph entail? It should be apparent at this point that formalism as I interpret it is not merely a conception of the common law or one of the proper role of the judge or of adjudication. Rather, it is a comprehensive program for law. It would therefore entail, if implemented, either that the restrained sense of ambition and competence I advocate be internalized both by judges and by legislators or that it be internalized by judges and (arrogantly!) employed by them to constrain legislators. If it is too late to return to *Lochner*, narrow interpretive strategies might be adopted.

It should be noticed that, while textualism is sometimes deemed a formalist strategy, it is not in fact clear whether it would enable or prevent a judiciary bent on constraining legislative excess. *Compare* SCALIA, *supra* note 149, at 29 (rejecting strict construction as anti-democratic and denying that textualism is anti-democratic), *with* Price Marshall, *No Political Truth: The Federalist and Justice Scalia on the Separation of Powers*, 12 U. ARK. LITTLE ROCK L. REV. 245, 253-54 (1989) (Scalia seeks to restrain legislature); David Schultz, *Judicial Review and Legislative Deference: The Political Process of Antonio Scalia*, 16 NOVA L. REV. 1249, 1265-71 (1992) (Scalia distrusts legislative process).

"absurdity" and "evil" are often employed without a sense of proportion and in service of utopian visions.

These "oftens" to one side, it remains the case that formalism demands tolerance of bad things, and under circumstances in which there is apparent power to correct them. This is not a tolerance much evident in contemporary value systems. The formalist's failure to correct apparent injustice has been denigrated as an escape from responsibility, evidence of adolescence, and as rendering the formalist himself the author of the evil he tolerates.[176] I think these characterizations unjustified, but they must be conceded to be popular.

Lest I be misunderstood, let me make it clear that I do not deny that great evils have been furthered by the law; although I think more great evils are associated with anti-formalism than with formalism. My points, rather, are that the distinction between great evils and unfortunate bads is not one much admired in contemporary America, that the resulting intolerance of unfortunate bads threatens formalism's empty spaces, and that this intolerance appears currently rampant.

Second, formalism isn't much fun, particularly from an intellectual point of view. I do not think formalism "easy" or unchallenging. Nor do I think the formalist in fact a mere automaton, applying without difficulty rule to fact. Both formalist rhetoric and anti-formalist rhetoric exaggerate formalism when they depict it as unproblematic rule following. Nevertheless, formalism is not unbridled moral philosophy, applied price theory or the ingenuous remaking of American society through the working out of a set of allegedly "preferred" values. It cannot, therefore, be attractive to persons with large intellectual ambitions. Law schools and the legal profession have for many years now attracted precisely such persons. The result is no doubt a vast improvement in the academic quality of the schools, and, perhaps, the intellectual power of the profession. I cannot help thinking that society would have been better off if this talent had applied itself within more socially productive fields, but this is not my point. My point is that formalism is not a likely candidate for fulfilling these ambitions.

In short, formalism, like other "isms," requires for its triumph compatibility with the self interest of the elites in a position to implement it. That condition is not satisfied.

---

176. FRANK, *supra* note 14. *Cf.* Alexander, *supra* note 47, at 562-64 (formalism as morally implausible).

ABUNDANT MEDIA, VIEWER SCARCITY: A MARKETPLACE ALTERNATIVE TO FIRST AMENDMENT BROADCAST RIGHTS AND THE REGULATION OF TELEVISED PRESIDENTIAL DEBATES

PAUL B. MATEY*

INTRODUCTION

The dramatic conclusion to the 2000 presidential election revealed a deeply divided nation. Voters split their choices throughout the country, sweeping out a host of Republican incumbents, while ending eight years of Democratic control in the White House. If the message sent to Washington was far from clear, so too was the motivation for the voters' choices. Although the candidates poured millions into commercials, Internet sites, and bus tours,[1] an undecided public focused on one campaign event: the debates. Despite the alternatives, and even with the inherent flaws, the live presidential debates became a singularly important source of information for American voters. Indeed, more than forty-six million households tuned in for the first debate, a number exceeding the first face-off between the candidates in 1996.[2]

The continued importance of the presidential debates, however, might soon prove insufficient to surmount the economics of network broadcasting. In the past decade, broadcast networks have watched viewers depart in record numbers, lured away by new technology and an ever-increasing array of media alternatives. With cable, satellite, and the Internet all vying for consumer attention, television networks continue to face slumping ratings and sagging profits.

Television networks have long questioned their role as guardians of the public interest. Today, with low-cost media alternatives eroding the television audience, the networks have launched a renewed attack on the Federal Communication Commission's policies on civic programming through court challenges, lobbying, and news editorials. In the 2000 election two networks, the National Broadcasting Company (NBC) and the FOX Network (FOX), stepped up their protests by simply refusing to air the first of the general election presidential debates. While sharp criticism rained down from the Commissioners, the ratings suggest that the American public welcomed the content choice. With no end to network troubles in sight, the 2004 elections

---

* Law Clerk to the Honorable Robert E. Cowen, United States Court of Appeals for the Third Circuit, 2002-03. J.D., Seton Hall University School of Law. I thank Mark Alexander, Matt Caudill, Joy Lindo, Scott Matey, and Dale Matey for their thoughtful suggestions and revisions. Comments may be directed to paul_matey@alum.shu.edu.

1. According to the Federal Election Committee, of the candidates who remained in the race through the election, Democratic presidential hopefuls spent $53,708,403, Republicans $94,466,341, and other parties a combined total of $14,428,180. Receipts of 1999-2000 Presidential Campaigns Through July 31, 2000, *available at* http://www.fec.gov/finance/precm8.htm (last visited Feb. 18, 2002). Candidates who withdrew before the election spent a combined total of $342,963,864. *Id.*

2. *See infra* note 213 and accompanying text.

could lack any unified television coverage.

Preserving the historic importance of the televised presidential debates will require the Federal Communications Commission (FCC) and the Supreme Court to confront the foundation of broadcast regulation in the United States. Historically, the Court has supported the FCC's duty to protect the public interest, relaxing the protections of the First Amendment to permit content-based restrictions on broadcasters. The basis for this public interest role, however, rests firmly on the doctrine of spectrum scarcity to justify the governmental grant of broadcast monopoly power. Scarcity theories, long criticized as economically inefficient, have now been attacked as scientifically flawed and incompatible with the new digital world. Today, many argue that increased competition in broadcast media provides a better guarantee of the public interest than intrusive government oversight. Trusting the market, however, might ignore the unequal access to new broadcast technologies and trap large numbers of the American electorate behind the digital divide.

A legislative solution is an attractive but unlikely answer. The FCC's public interest power, which includes authority to regulate broadcast indecency, is difficult for politicians to attack directly without loss of political capital. Moreover, the 2000 presidential elections highlighted more pressing deficiencies in the voting process, problems that remain in the national spotlight.[3] Administrative options within the FCC are equally unlikely, given its size, structure, and partisan composition.

Judicial intervention, sometimes criticized in other areas of national debate, is the best solution. The Supreme Court holds the unique responsibility of defining the First Amendment rights of broadcasters. The Court is both the historic arbiter of the Constitution and the modern source of the FCC's sweeping regulatory authority. A solution to the chaos of the First Amendment rights of broadcasters is necessary and available in the same economic analysis that supports the criticism of the current state of the law. This Article suggests that the First Amendment rights of broadcasters should be evaluated using the market power of the broadcast content to determine the degree of constitutional protection. This Article then applies this new standard of review to a model broadcast debate regulation, which compels the major television networks to provide live coverage of the general presidential debates.

Part I of this Article recounts the history of American broadcast regulation. Tracing the development of the FCC through statute and commentary, Part I outlines the doctrines of scarcity that underlie the FCC's public interest mission. Noting the economic irrationality of the scarcity theory and its conflict with First Amendment values, Part I concludes that scarcity does not justify continued federal oversight. Part II continues with a discussion of the past and present importance of live televised debates in the general presidential election.[4] This

---

3. Edward Walsh & Dan Balz, *One Year Later, Election Reform Remains Elusive*, WASH. POST, Nov. 13, 2001, at A3.

4. Broadcast coverage of political debates has raised concerns outside the general presidential elections. Commentators have addressed problems regarding coverage of the

section also explains the 2000 presidential debates and the decision of NBC and FOX to decline live coverage.

With this background in mind, Part III offers a fresh solution based on the reasoning of the Supreme Court's two most recent First Amendment broadcasting decisions, *Reno v. ACLU*[5] and *Turner Broadcasting System, Inc. v. FCC* (*Turner I*).[6] After an overview of the Court's content approach to speech, Part III explains the Court's forum-specific approach to the First Amendment in broadcasting. The Court's attention to the converging markets for broadcast speech suggests a finite future for the scarcity doctrine and a new technology-specific approach to the First Amendment in broadcasting. Part III then explains an alternative market-based approach to the First Amendment rights of broadcasters using the product and geographic market standards developed in antitrust economics. Market power, Part III argues, provides a dividing line between the scarce media of broadcast television and radio and the plentiful resources of the digital spectrum. Full First Amendment protections for only converged[7] broadcast media, Part III concludes, will retain administrative regulation over lagging technologies while inducing broadcasters to speed the development of broadband.

Part IV constructs a model regulation compelling the coverage of the general presidential debates. After outlining the suggested goals of a debate rule, Part IV tests the model against the market-based First Amendment review. Despite the emerging alternatives to broadcast television and the competing sources of campaign information, Part IV concludes that national televised coverage of the general presidential debates comprises a single, powerful content market. Technological advancements and a national spirit of campaign reform could soon transform our understanding of broadcast political coverage. Until then, this

---

presidential primaries, and contests for seats in both houses of Congress. *See, e.g.*, Jamin B. Raskin, *The Debate Gerrymander*, 77 TEX. L. REV. 1943 (1999). Others have noted the problems posed by the Supreme Court's decision in *Arkansas Education Television Commission v. Forbes*, 523 U.S. 666 (1998), which held that states may exclude so-called "third-party candidates" on neutral criteria such as the "public interest" in the candidate. *See, e.g.*, Kyu Ho Youm, *Editorial Rights of Public Broadcasting Stations vs. Access for Minor Political Candidates to Television Debates*, 52 FED. COMM. L.J. 687 (2000); Keith Darren Eisner, Comment, *Non-Major-Party Candidates and Televised Presidential Debates: The Merits of Legislative Inclusion*, 141 U. PA. L. REV. 973 (1993). The general presidential election debates, however, "occur in the one campaign that commands considerable voter attention." Daniel H. Lowenstein, Commentary, *Election Law Miscellany: Enforcement, Access to Debates, Qualification of Initiatives*, 77 TEX. L. REV. 2001, 2010 (1999). Accordingly, there might be "no basis in evidence or in common sense for extrapolating from the importance of presidential general election debates a similar significance" to other debates. *Id.* Without addressing the merits of the comparison, this Article assumes that the general presidential election debates are indeed "major events," different in scope and meaning from other debates. *Id.* The analysis of this Article is limited accordingly.

5. 521 U.S. 844 (1997).

6. 512 U.S. 622 (1994).

7. *See infra* notes 247-48 and accompanying text.

Article suggests that federal oversight of the televised presidential debates is an appropriate and necessary limitation on the First Amendment rights of networks.

I. FEDERAL REGULATION OF THE AIRWAVES: A BRIEF HISTORICAL OVERVIEW OF DUTY AND THEORY

Broadcast regulation emerged from a combination of military pressure, international tragedy, and a limited understanding of technology. Given a broad statutory mandate by Congress, federal broadcast regulation sought to order the multitude of speakers rushing to the newly discovered airwaves.[8] Throughout the Twentieth Century, the modest regulatory goal initially conceived by Congress developed a wide social mission through federal regulations aimed at improving the quality of public debate.[9] During the same period, the analytic foundations for these concededly laudatory goals have been continually questioned and repeatedly marginalized. An examination of these foundations, based on the doctrine of scarcity, illustrates that additional control over broadcasting could strain First Amendment precedents despite the importance of preserving an informed electorate.

A. The Early History: Ships, Radios, and Secretary Hoover

Federal control of the American broadcast airwaves emanated from the tragedy of the *Titanic*.[10] In the early 1900s, broadcast radio emerged as a commercial force in naval communications, primarily in the business of private shipping.[11] The *Titanic* disaster highlighted the growing concerns of the U.S. Navy that autonomous, unregulated radio broadcasters impeded the safe passage of military and commercial vessels, creating a state of chaos on the seas.[12]

---

8. *See* Marc Sophos, Comment, *The Public Interest, Convenience, or Necessity: A Dead Standard in the Era of Broadcast Deregulation?*, 10 PACE L. REV. 661, 666 (1990).

9. *See* LEE C. BOLLINGER, IMAGES OF A FREE PRESS 65 (1991).

10. *See* THOMAS G. KRATTENMAKER & LUCAS A. POWE, JR., REGULATING BROADCAST PROGRAMMING 5 (1994). Although the *Titanic* sent several distress calls in the hours before the ship disappeared, "amateur radio operators along the East Coast filled the air with questions, rumors, and, most of all, interference." *Id.*

11. *See* R.H. Coase, *The Federal Communications Commission*, 2 J.L. & ECON. 1, 1-2 (1959). The growth of maritime commerce led to the passage of the Wireless Ship Act of 1910, which required ships leaving a United States port with more than fifty passengers to be equipped with radio transmitters. Wireless Ship Act of June 24, 1910, ch. 379, 36 Stat. 629.

12. *See* S. REP. NO. 659, 61st Cong. § 3 (1912), *reprinted in* Coase, *supra* note 11, at 2. The Navy further claimed that "[c]alls of distress from vessels in peril on the sea go unheeded or are drowned out . . . ." Coase, *supra* note 11, at 2. In addition, the ratification of the first international radio treaty in 1912 required the United States to develop uniform policies of broadcasting that were compatible with international use. *See* ROBERT L. HILLIARD, THE FEDERAL COMMUNICATIONS COMMISSION: A PRIMER 61 (1991). The combination of military pressure and international obligation moved Congress to action. *See* Mike Harrington, *A-B-C, See You Real Soon: Broadcast Media Mergers and Ensuring a "Diversity of Voices,"* 38 B.C. L. REV. 497, 503 (1997).

Congress responded by seizing the radio airwaves and requiring radio operators to seek licenses from the Department of Commerce. Spurred by the Navy, the Radio Act of 1912[13] created a regulatory system that favored large-scale commerce,[14] and not surprisingly, military defense.[15]

The 1912 Act, however, proved insufficient to control the rapid growth of the radio industry in the 1920s[16] as private operators rushed to develop the new market.[17] Seeking uniformity, then Secretary of Commerce Herbert Hoover began denying new commercial radio licenses.[18] Soon after, Hoover's power to condition licenses under the 1912 Act was removed by court challenge,[19] leaving the radio industry without effective federal oversight. As the Supreme Court later observed, the result was again "chaos."[20] Following a showdown between Secretary Hoover and the broadcasting industry,[21] Congress passed the Radio Act

---

13. The Radio Act of 1912, Pub. L. No. 62-264, 37 Stat. 302.

14. *See* KRATTENMAKER & POWE, *supra* note 10, at 6. The authors note that while ships were granted an exclusive portion of the broadcast spectrum, amateur operators were "relegated to oblivion." *Id. See also* Coase, *supra* note 11, at 3 (noting that under the 1912 Act, amateur broadcasters were limited to wavelengths less than two hundred meters).

15. *See* KRATTENMAKER & POWE, *supra* note 10, at 6. In particular, the 1912 Act allowed the military to seize all radio signals and equipment in wartime. *Id.*

16. Thomas W. Hazlett, *The Rationality of U.S. Regulation of the Broadcast Spectrum*, 33 J.L. & ECON. 133, 139 (1990) (reporting that by 1922 more than 576 broadcast stations were transmitting in the United States).

17. Karen Beth Gray, Note, *Fairness Doctrine Termination: Extinction of an Unenforceable Theory*, 22 SUFFOLK U. L. REV. 1057, 1058-59 (1988).

18. *See* KRATTENMAKER & POWE, *supra* note 10, at 9. Secretary Hoover had originally sought to broker wider industry regulation through a series of meetings between government agencies and commercial broadcasters called the National Radio Conferences. *See* Coase, *supra* note 11, at 4. Although the meetings produced a series of legislative recommendations, Congress adopted none of the proposals. *See id.*

19. Hoover v. Intercity Radio, Inc., 286 F. 1003 (D.C. Cir. 1923). Secretary Hoover sought to limit the number of successful licensees by drafting detailed conditions into the applications. *See* Coase, *supra* note 11, at 4. In 1923, Secretary Hoover convened another meeting of the National Radio Conference, which concluded that the 1912 Act permitted the Secretary to regulate both the frequencies of radio broadcasts, and the hours of operation for radio licensees. *See* Red Lion Broad. Co. v. FCC, 395 U.S. 367, 375 n.4 (1969). Soon after, however, the Court of Appeals for the District of Columbia held that the Secretary's role in the licensing process under the 1912 Act was limited to the selection of the wavelength for the applicant. *See Intercity Radio*, 286 F. at 1007. Accordingly, the Act of 1912 reposed "no discretion whatever in the Secretary of Commerce," and made the issuance of a license "mandatory." *Id.*

20. *Red Lion*, 395 U.S. at 375.

21. *See* Hazlett, *supra* note 16, at 141. Although the *Intercity Radio* decision had limited the regulatory power of the Commerce Department, Secretary Hoover began refusing to process new applications in defiance of the court's order. *See id.* Hoover's actions were again disapproved by court decision, and his powers limited once more to only "the regulations in the Act itself." Coase, *supra* note 11, at 5 (discussing United States v. Zenith Radio Corp., 12 F.2d 614 (N.D. Ill. 1926));

of 1927[22] and created a new agency, the Federal Radio Commission (FRC).[23] The 1927 Act marked the beginning of serious federal communications oversight.[24]

In 1934, Congress tightened the regulation of broadcasting by replacing the FRC with the FCC.[25] Congress gave the new seven-member commission[26] wide regulatory power over all broadcast media, including radio, telegraph, and telephone.[27] Congress also carried over the FRC's basic mandate into the 1934 Act, empowering the FCC to issue broadcast licenses[28] for the "public interest, convenience, or necessity."[29]

B. The Public Interest Standard Explained and Applied

The public interest standards of the Acts of 1927 and 1934 provided the FCC with general authority to protect the public interest.[30] The public interest doctrine was not an entirely new concept; it had been used elsewhere in federal legislation governing state-created monopolies[31] and private control of public

---

see also Fed. Regulation of Radio Broad., 35 Op. Att'y Gen. No. 126, 129 (1926) (agreeing with the *Zenith* court's interpretation of the 1912 Act, and concluding that "[t]he power to make general regulations is nowhere granted by specific language to the Secretary").

Rather than continuing to challenge the regulatory limits imposed by the 1912 Act, Secretary Hoover "issued a statement abandoning all his efforts to regulate radio and urging that the stations undertake self-regulation." NBC v. United States, 319 U.S. 190, 212 (1943). With federal oversight removed, a flood of new broadcasters swarmed the airwaves, and "[m]ore than two hundred stations were established in the next nine months." Coase, *supra* note 11, at 5.

22. Radio Act of 1927, ch. 169, 44 Stat. 1162.

23. *Id.* § 3, at 1162.

24. *See* Timothy B. Dyk & Ralph E. Goldberg, *The First Amendment and Congressional Investigations of Broadcast Programming*, 3 J.L. & POL. 625, 628 (1987).

25. *See* Communications Act of 1934, ch. 652, 48 Stat. 1064.

26. *Id.* § 4(a), at 1066. In 1986, Congress reduced the FCC's membership to five. Pub. L. No. 99-334, 100 Stat. 513, 47 U.S.C. § 154(a).

27. *See* DONALD J. JUNG, THE FEDERAL COMMUNICATIONS COMMISSION, THE BROADCAST INDUSTRY, AND THE FAIRNESS DOCTRINE: 1981-1987, at 8 (1996). The 1934 Act, however, retained substantially all of the regulatory framework of the 1927 Act. *See* FCC v. Pottsville Broad. Co., 309 U.S. 134, 137 (1940). The FCC was permitted to issue licenses for three-year periods, where the license would benefit the public interest. Communications Act of 1934, ch. 652, § 307, 48 Stat. 1064, at 1083-84.

28. For an overview of the modern broadcast licensing process, see Timothy B. Dyk, *Full First Amendment Freedom for Broadcasters: The Industry as Eliza on the Ice and Congress as the Friendly Overseer*, 5 YALE J. ON REG. 299, 301-02 (1988).

29. Communications Act of 1934, ch. 652, § 602(a), 48 Stat. 1064.

30. *See* BOLLINGER, *supra* note 9, at 63 (describing the public interest doctrine in broadcasting as "the most general mandate imaginable").

31. *See, e.g.*, United Shoe Mach. Corp. v. United States, 258 U.S. 451, 464 (1922) (holding that Section Three of the Clayton Act is not an unconstitutional restriction on the rights of patent

resources.[32] The Supreme Court first explained the public interest standard of the Communications Acts in *NBC v. United States*, stating that the public interest doctrine assumes that broadcast regulation should "secure the maximum benefits . . . to all the people of the United States."[33] Recounting the chaotic results that followed the narrow interpretations of the 1912 Act,[34] the Court concluded that Congress had premised the Act of 1927 on the belief that federal regulation was essential to avoid wasting the broadcast airwaves.[35] The 1927 Act thus established a "unified and comprehensive regulatory system" to manage broadcast traffic.[36]

The Court added, however, that the 1927 Act "does not restrict the Commission merely to supervision of the traffic."[37] The Court noted that the "dynamic" nature of radio necessitated broad legislative language capable of evolving with new developments in the broadcast medium.[38] The 1927 Act, according to the Court, delegated to the FCC the task of determining the "larger and more effective use of radio."[39] The Court supported this broad interpretation by citing the physical limits of the broadcast spectrum, reasoning that because "[t]he facilities of radio are not large enough to accommodate all who wish to use them," the FCC is authorized to determine "the composition of [the] traffic."[40] The Court thus used the chaos of unregulated broadcasting[41] to justify both procedural and substantive regulation of the airwaves.[42]

---

holders, as Congress may prohibit "in the public interest the making of agreements which may lessen competition and build up monopoly").

32. *See, e.g.*, Transportation Act of 1920, Pub. L. No. 66-152, 41 Stat. 456, 477-78 (requiring the Interstate Commerce Commission to determine whether a proposed extension to a railroad is required for the "present or future public convenience and necessity"); Bd. of Trade of Chicago v. Olsen, 262 U.S. 1, 41 (1923) (holding that because the Chicago grain exchange is a business "affected with a public national interest" it is "subject to national regulation as such"); Munn v. Illinois, 94 U.S. 113, 126 (1876) (holding that when private property is devoted to a public interest, the owner "must submit to be controlled by the public for the common good, to the extent of the interest . . . created").

33. NBC v. United States, 319 U.S. 190, 217 (1943).

34. *See supra* notes 13-21 and accompanying text.

35. *NBC*, 319 U.S. at 213.

36. *Id.* at 214. The Court specifically noted that "[r]egulation of radio was . . . as vital to its development as traffic control was to the development of the automobile." *Id.* at 213.

37. *Id.* at 215-16.

38. *Id.* at 219-20.

39. *Id.* at 216. The Court viewed the 1927 Act as an unremarkable exercise of legislative delegation, defining "broad areas for regulation" and general "standards for judgment." *Id.* at 219-20.

40. *Id.* at 215-16.

41. *See* FCC v. Sanders Bros. Radio Station, 309 U.S. 470, 475 (1940).

42. In *NBC*, the Court stated that the 1934 Act is not solely concerned with the regulation of broadcast traffic, but also "puts upon the Commission the burden of determining the composition of that traffic." *NBC*, 319 U.S. at 215-16. In *CBS v. Democratic National Committee*, 412 U.S.

One proposition, however, does not necessarily lead to the next.[43] Congress had passed the 1927 Act to correct the market failures in broadcasting and had incorporated the public interest standard to avoid the restrictive judicial interpretations that led to the chaos of the early 1900s.[44] Neither goal necessarily required the FCC to become the public guardian of broadcast content. Despite this potential flaw, broader applications of the public interest doctrine[45] have continued to rely on the Court's reasoning in *NBC* and the Court's expanded discussion of the doctrine in *Red Lion Broadcasting v. FCC*.[46] This reasoning, described under the catchphrase "scarcity," remains the foundation for federal broadcast oversight[47] and thus must be considered as a likely source of authority for regulation of the presidential debates.

C. The Many Faces of Scarcity

More than fifty years ago the Supreme Court held that broadcasting enjoyed the protections of the First Amendment.[48] Despite this fact, Americans have comfortably accepted pervasive regulation of broadcasting, regulations that

---

94, 117 (1973), the Court further noted that although broadcast regulation had developed slowly, the FCC now acts as an "'overseer' and ultimate arbiter and guardian of the public interest." The Court has therefore interpreted the public interest doctrine as a "supple instrument for the exercise of discretion by the expert body." FCC v. Pottsville Broad. Co., 309 U.S. 134, 138 (1940). Ensuring that broadcasting serves the public interest, moreover, might allow the Commission to rest on "judgment and prediction" rather than pure factual determinations. FCC v. Nat'l Citizens Comm. for Broad., 436 U.S. 775, 814 (1978). Decisions of the FCC "regarding how the public interest is best served [are] entitled to substantial judicial deference." FCC v. WNCN Listeners Guild, 450 U.S. 582, 596 (1981).

43. *See, e.g.*, Louis L. Jaffe, *The Illusion of the Ideal Administration*, 86 HARV. L. REV. 1183, 1191-92 (1973). Professor Jaffe argued that the 1927 Act sought only to address the judicial decisions prohibiting the Secretary of Commerce from limiting licenses in order to manage airwave traffic: "[T]he use of 'public interest' in the statute did not manifest a congressional intent to give the Commission general powers to 'regulate' the industry or to solve any 'problem' other than the problem of interference . . . ." *Id.*

44. *See* Erwin G. Krasnow & Jack N. Goodman, *The "Public Interest" Standard: The Search for the Holy Grail*, 50 FED. COMM. L.J. 605, 609-10 (1998). *See also supra* notes 19-21 and accompanying text.

45. *See* Stephen F. Varholy, *Preserving the Public Interest: A Topical Analysis of Cable/DBS Crossownership in the Rulemaking for the Direct Broadcast Satellite Service*, 7 COMM. L. CONSPECTUS 173, 175-76 (1999) (describing the FCC's public interest duties as diversity of content and competition among broadcasters); *see also* William T. Mayton, *The Illegitimacy of the Public Interest Standard at the FCC*, 38 EMORY L.J. 715, 716 (1989). Professor Mayton argues that the FCC developed its broad public interest mission not from statute, but from an internal agency belief that "progressive social change is best accomplished through government regulation . . . ." *Id.* at 717.

46. 395 U.S. 367 (1969).

47. Turner Broad. Sys., Inc. v. FCC, 512 U.S. 622, 638-39 (1994).

48. *See* United States v. Paramount Pictures, Inc., 334 U.S. 131, 166 (1948).

would seem inappropriate in newsprint, books, or sidewalk speech.[49] The Court has upheld these regulations by creating a two-tiered system of First Amendment analysis. First, the Court differentiates between the protection afforded to broadcasting and physical media, and second, it varies the protection afforded within the broadcast media, including radio, network television, cable television, and the Internet. The Supreme Court's decisions rely heavily on assumptions regarding the physical aspects of broadcasting and the scarcity of electromagnetic space.

The scarcity doctrine is a seemingly simple concept, with origins in common sense, if not science. The chaos of early radio broadcasting stemmed from too many users and too few frequencies. Order was imposed by the 1927 Act, which allocated the broadcast spectrum through a licensing system that explicitly reserved the ownership of the airwaves for the public.[50] Broadcasters would have a mere right of access based on their willingness or ability to serve the public interest.[51] At the same time, however, broadcasting was and is speech protected by the First Amendment.[52] Limiting broadcast speech for orderly use or social gain would thus seem to conflict with constitutional protections.[53] Scarcity, therefore, became the necessary analytical "problem" to justify broadcast restraints.[54]

The scarcity doctrine originated in *NBC*, where Justice Frankfurter described "certain basic facts" about radio broadcasting: "its facilities are limited," and "the radio spectrum simply is not large enough to accommodate everybody. There is a fixed natural limitation upon the number of stations that can operate . . . ."[55] The Court viewed this natural limitation as "unique," distinguishing broadcasting from other forms of speech.[56] Accordingly, government regulation was necessary to select which of the many speakers seeking access to the

---

49. *See* BOLLINGER, *supra* note 9, at 62.

50. 47 U.S.C. § 304 (1988). The statute states that

> No station license shall be granted by the Commission until the applicant therefor shall have waived any claim to the use of any particular frequency or of the electromagnetic spectrum as against the regulatory power of the United States because of the previous use of the same, whether by license or otherwise.

Id.

51. *See* Reed E. Hundt, *A New Paradigm for Broadcast Regulation*, 15 J.L. & COM. 527, 528-29 (1996).

52. *Paramount Pictures*, 334 U.S. at 166 (holding there is "no doubt that . . . newspapers and radio . . . are included in the press whose freedom is guaranteed by the First Amendment").

53. *See* Joseph Burstyn, Inc. v. Wilson, 343 U.S. 495, 502-03 (1952). Although the Court has not held that every method of communication is "necessarily subject to the precise rules governing any other particular method of expression," it has also noted "the basic principles of freedom of speech and the press, like the First Amendment's command, do not vary." *Id.*

54. *Id.* at 503 ("Each method [of speech] tends to present its own peculiar problems.").

55. NBC v. United States, 319 U.S. 190, 213 (1943).

56. *Id.* at 226.

airwaves should be admitted.[57]

The Supreme Court reiterated its position on scarcity in *Red Lion*[58] and increased its reliance on government oversight as the only means of regulating broadcasters. Again citing the limits of the broadcast spectrum,[59] the Court found it "essential" that regulation allocate the airwaves among competing speakers.[60] Moreover, the Court held that *only* control by the federal government was sufficient.[61] Whereas *NBC* concluded that some sort of regulation of the airwaves was necessary[62] and that federal regulation of the airwaves was constitutional, *Red Lion* stated that no alternatives to governmental control were possible.[63]

Courts[64] and commentators[65] have criticized the scarcity doctrines[66] articulated in *NBC* and *Red Lion* as inconsistent with basic principles of a free-market economy. First, as Professor Coase observed in 1959, the mere scarcity of an important resource does not normally justify government regulation.[67] The

---

57. *See id.* at 216 (holding that because the "facilities of radio are not large enough to accommodate all who wish to use them . . . [m]ethods must be devised for choosing from among the many who apply").

58. Red Lion Broad. Co. v. FCC, 395 U.S. 367, 375 (1969).

59. *See id.* at 388.

60. *Id.*

61. *See id.* at 376.

62. *See NBC*, 319 U.S. at 213.

63. *See Red Lion*, 395 U.S. at 376 ("Without government control, the medium would be of little use because of the cacophony of competing voices, none of which could be clearly and predictably heard.").

64. *See, e.g.*, Action for Children's Television v. FCC, 58 F.3d 654, 673-76 (D.C. Cir. 1995) (en banc) (Edwards, C.J., dissenting); Telecomm. Research & Action Ctr. v. FCC, 801 F.2d 501, 508 (D.C. Cir. 1986).

65. *See, e.g.*, Coase, *supra* note 11; KRATTENMAKER & POWE, *supra* note 10; Charles W. Logan, Jr., *Getting Beyond Scarcity: A New Paradigm for Assessing the Constitutionality of Broadcast Regulation*, 85 CAL. L. REV. 1687 (1997); Murray J. Rossini, *The Spectrum Scarcity Doctrine: A Constitutional Anachronism*, 39 SW. L.J. 827 (1985).

66. For a comprehensive discussion of the various forms of the scarcity rationale, see KRATTENMAKER & POWE, *supra* note 10, at 204-19, and Matthew L. Spitzer, *Controlling the Content of Print and Broadcast*, 58 S. CAL. L. REV. 1351, 1358-64 (1985).

67. Coase, *supra* note 11, at 14. Professor Coase summarized the economic criticism of the scarcity doctrine:

> [I]t is a commonplace of economics that almost all resources used in the economic system (and not simply radio and television frequencies) are limited in amount and scarce, in that people would like to use more than exists. Land, labor, and capital are all scarce, but this, of itself, does not call for government regulation. It is true that some mechanism has to be employed to decide who, out of the many claimants, should be allowed to use the scarce resource. But the way this is usually done in the American economic system is to employ the price mechanism, and this allocates resources to users without the need for government regulation.

"chaos" throughout broadcasting prior to 1927 occurred because of an absence of property rights in the broadcast spectrum.[68] The marketplace, in turn, can allocate property rights without the government oversight condoned under *NBC* or thought mandatory under *Red Lion*.[69]

Second, the entry barriers to broadcasting caused by the limits to the electromagnetic spectrum are present in analogous media, such as newsprint.[70] The Court seemingly agreed with this conclusion in *Miami Herald Publishing Co. v. Tornillo*.[71] *Tornillo* involved a state statute requiring newspapers to print editorial replies from candidates personally or professionally assailed in the same paper.[72] The Court observed that the scarcity of newspapers and the costs of starting an independent publication created an entry barrier sufficient to silence the speech of persons denied access to established papers.[73] Nonetheless, the scarcity of newspapers did not sway the Court's conclusion that the First Amendment prevents governmental regulation of publishers[74] and forbids restrictions designed to foster a "responsible press."[75] The Court's acknowledgment of the limited resources intrinsic to both newspapers and broadcasting makes scarcity a tenuous ground for limiting the First Amendment rights of broadcasters.[76]

Third, the physical limitations on the broadcast airwaves cited in both *NBC* and *Red Lion* might no longer exist.[77] Throughout the Twentieth Century, communications technologies began to travel beyond the electromagnetic spectrum. In 1950, cable reached only 14,000 televisions in America,[78] a figure that would rise to more than sixty-five million by 1998.[79] Direct broadcast

---

Id.

68. KRATTENMAKER & POWE, *supra* note 10, at 207.

69. Spitzer, *supra* note 66, at 1360-61.

70. *See, e.g.*, WILLIAM W. VAN ALSTYNE, FIRST AMENDMENT: CASES AND MATERIALS 546-47 (1995).

71. 418 U.S. 241 (1974).

72. *Id.* at 244.

73. *Id.* at 251.

74. *Id.* at 258.

75. *Id.* at 256 (concluding that while a "responsible press is an undoubtedly desirable goal" it is "not mandated by the Constitution and like many other virtues it cannot be legislated").

76. *See* Telecomm. Research & Action Ctr. v. FCC, 801 F.2d 501, 508-09 (D.C. Cir. 1986) (discussing the scarcity of physical media such as "newsprint, ink, delivery trucks, computers, and other resources that go into the production . . . of print journalism").

77. Indeed, as Professor Thomas Hazlett has documented, physical spectrum scarcity may never have existed. *See* Thomas W. Hazlett, *Physical Scarcity, Rent Seeking, and the First Amendment*, 97 COLUM. L. REV. 905, 926-31 (1997). Professor Hazlett noted that radio programming was delivered via cable by 1923, suggesting that technological alternatives to the electromagnetic spectrum existed concurrently with the emergence of broadcasting. *Id.* at 928-29.

78. FCC, FACT SHEET: GENERAL INFORMATION, CABLE TV AND ITS REGULATION (June 2000), *available at* http://www.fcc.gov/mb/facts.

79. *Id.*

satellites (DBS) are now installed in more than eighteen million homes, an increase of approximately two million subscribers since 2001.[80] Satellites offer radio listeners a similar array of programming choices without the regional limitations inherent in traditional radio broadcasting.[81] Internet access now reaches an estimated fifty-four million American subscribers[82] with 143 million people, or more than fifty-three percent of the population, using the Internet.[83]

Formerly distinct industries such as telephony have converged[84] with broadcasting to offer new forms of digital programming. Convergence is leading to the growth of interactive television services[85] such as video-on-demand, email, gaming, and electronic commerce.[86] In addition, several cable and satellite providers[87] offer broadband technologies[88] capable of transmitting graphics,

---

80. *In re* Annual Assessment of the Status of Competition in the Market for the Delivery of Video Programming, FCC 02-338 (2002), 2002 WL 31890210 para. 58 [hereinafter *In re* Annual Assessment].

81. Amanda Barnett, *Radio About to Go Higher Tech*, *at* http://www.cnn.com (May 23, 2001). Satellite radio networks offer more than one hundred channels of programming, including news from companies such as FOX, National Public Radio, AP Radio, the BBC, and C-Span. *Id.* The first entrant into the market, XM Satellite Radio, has sold more than 25,000 subscriptions since its debut in November 2001, a record among new audio products in recent years. Laurie J. Flynn, *Satellite Radio Shows Growth*, N.Y. TIMES, Jan. 7. 2002, at C7. Industry analysts estimate that as many as twenty-five million people will subscribe to satellite radio by 2009. David Becker, *A Satellite Radio Field of Dreams*, CNET News (June 8, 2002), *at* http://news.com.com/2100-1033-803900.html.

82. *In re* Annual Assessment, *supra* note 80, para. 89.

83. ECONOMICS AND STATISTICS ADMINISTRATION & NATIONAL TELECOMMUNICATIONS AND INFORMATION ADMINISTRATION, A NATION ONLINE: HOW AMERICANS ARE EXPANDING THEIR USE OF THE INTERNET 10 (2002) [hereinafter ESA, NATION ONLINE], *available at* http://www.ntia.doc.gov/ntiahome/dn/index.html.

84. As commonly understood, convergence "refers to the coming together of several formerly distinct services and industries . . . into a single, digital marketplace." William T. Lake et al., *Telecommunications Convergence*, *in* TELECOMMUNICATIONS CONVERGENCE: IMPLICATIONS FOR THE INDUSTRY AND FOR THE PRACTICING LAWYER 11 (2000). The Supreme Court discussed the importance of convergence in *Turner Broadcasting Systems, Inc. v. FCC*, 512 U.S. 622, 627 (1994), noting "convergence between cable and other electronic media . . . [places] the cable industry . . . at the center of an ongoing telecommunications revolution . . . ."

85. *In re* Annual Assessment, *supra* note 80, para. 170. The FCC defines interactive television as "a service that supports subscriber-initiated choices or actions that are related to one or more video programming streams." *In re* Nondiscrimination in the Distribution of Interactive Television Services Over Cable, 16 F.C.C.R. 1321, 1323 (2001) [hereinafter *In re* Interactive Television].

86. *In re* Annual Assessment, *supra* note 80, para. 170; *see also In re* Interactive Television, *supra* note 85, at 1323-28.

87. *Id.* (describing interactive programming, including high-speed Internet access and interactive television).

88. Karen Kornbluh, Editorial, *The Broadband Economy*, N.Y. TIMES, Dec. 10, 2001, at A21

video, and data at more than four times the speed of dial-up telephone modems.[89] Convergence is now more than technological theory, with pundits and executives united in praising the economic and social promise of a single media pipeline.[90]

Finally, local broadcasters have demonstrated new uses for the existing electromagnetic spectrum. Low-power radio frequencies capable of reaching one to two miles are now up for auction,[91] and low-power television offers local access outside of metropolitan areas.[92]

These rival technologies challenge the basis of the scarcity doctrine.[93] If

---

(defining broadband as "the generic term for high-speed, high-capacity, always-on data networks").

89. *In re* Inquiry Concerning the Deployment of Advanced Telecomm. Capability to All Americans in a Reasonable and Timely Fashion, and Possible Steps to Accelerate Such Development Pursuant to Section 706 of the Telecommunications Act of 1996, 15 F.C.C.R. 20,913, 20,920 (2000) [hereinafter *In re* Deployment]. The FCC defines high-speed Internet access as the "capability of supporting, in both the provider-to-customer (downstream) and the customer-to-provider (upstream) directions, a speed . . . in excess of 200 kilobits per second (kbps) in the last mile." *Id.* However, high-speed Internet access remains limited with just over fourteen million broadband subscribers as of June 2002. *In re* Annual Assessment, *supra* note 80, para. 88.

90. *See, e.g.*, Seth Schiesel, *AOL Plans the Digital Smorgasbord*, N.Y. TIMES, June 11, 2001, at C1. Steve Case, then chairman of AOL Time Warner, predicted that convergence will knit together "the PC, the TV, the telephone and the stereo to allow people to be entertained in better ways, to be educated in better ways, to communicate in better ways, to change people's lives." *Id.*; *see also* PETER HUBER, LAW AND DISORDER IN CYBERSPACE: ABOLISH THE FCC AND LET COMMON LAW RULE THE TELECOSM 23 (1997) ("The telecosm is being transformed into a network of networks, an intricately interconnected matrix of wireless, [and] satellite . . . with multiple overlapping and complementary providers, and no single dominant center."). Nonetheless, convergence is still in its infancy and vulnerable to continuing setbacks in the marketplace. Susan Stellin, *A Device to Link Old Media to the Web Struggles to Make Good on the Promise of an Internet Revolution*, N.Y. TIMES, Jan. 15, 2001, at C4 (noting that "one of the biggest challenges that has always stood in the way of convergence is the need to persuade so many different participants to mold their behavior or business strategy to an unknown technology").

Moreover, the American media consumer has shown a marked disinterest in some of the earliest, and most promoted, forms of convergence. Microsoft's highly touted WebTV, which permits television viewers to navigate the Internet over television screens, proved a commercial disappointment. Saul Hansell, *Clicking Outside the Box*, N.Y. TIMES, Sept. 20, 2000, at H1. One industry executive, Michael Willner, then president of Insight Communications, blamed the failure on the users, lamenting that "[p]eople want their information spoon-fed to them," and will therefore not embrace technologies requiring an active television viewer. *Id.* Others have suggested that the problem rests not with the audience, but with the programmers, arguing that "[c]onsumer[s] are slow to adopt broadband because, while there may be an infinite number of channels, there is still nothing on." Lawrence Lessig, *Who's Holding Back Broadband?*, N.Y. TIMES, Jan. 8, 2002, at A17.

91. FCC, Major Initiatives, *at* http://www.fcc.gov/major.html (last visited Mar. 30, 2002).

92. PRACTICING LAW INSTITUTE, NEW PROGRAM OPPORTUNITIES IN THE ELECTRONIC MEDIA (George H. Shapiro ed., 1983).

93. An alternative to the scarcity doctrine—sometimes known as the prior grant theory—has

wired or wireless signals can circumvent any limitations in the electromagnetic spectrum,[94] broadcasting may shed its unique status within the First Amendment. Despite the modern advances to date, however, the scarcity doctrine has proven surprisingly resilient in the Supreme Court.[95] The Court has consistently held that the physical scarcity of the broadcast airwaves underlies the Radio Act of 1927 and the Communications Act of 1934[96] and justifies a less rigorous degree of scrutiny than otherwise demanded by the First Amendment.[97] Specifically,

---

been equally criticized. This theory, first articulated by the Supreme Court in *Red Lion*, holds that broadcasters enjoy their market position through "a preferred position conferred by the Government," or a governmentally created monopoly. Red Lion Broad. Co. v. FCC, 395 U.S. 367, 400 (1969). The prior grant theory assumes that subjecting broadcasters to regulatory oversight is a suitable trade-off for the benefits of their dominant market position. *Id.* at 391. The prior grant theory, however, fails to explain the continuing role of federal oversight over broadcast industries that now enjoy vigorous competition. Mark S. Fowler & Daniel L. Brenner, *A Marketplace Approach to Broadcast Regulation*, 60 TEX. L. REV. 207, 226-27 (1982). Moreover, as noted by Professors Krattenmaker and Powe, the prior grant doctrine "proves too much," seemingly justifying a suspension of constitutional protections on any public forum in which the government claims ownership. KRATTENMAKER & POWE, *supra* note 10, at 228.

94. Hazlett, *supra* note 77, at 929. Professor Hazlett notes bluntly:

> The ability to substitute wired frequencies for wireless spectrum space should be self-evident today, when consumers and businesses choose daily between the rival forms of communications transmissions—for example, when deciding whether to use a TV antenna or satellite dish versus a cable TV hook-up, or placing a telephone call via a landline versus a cellphone (or cordless phone).

Id. Professor Hazlett's observation is supported by recent developments in the telecommunications industry, where the increased demand for spectrum space caused by wireless technologies has led to new ways to "increase the capacity and the efficiency of the available spectrum." Editorial, *Space Invaders*, WALL ST. J., June 5, 2001, at A26.

95. KRATTENMAKER & POWE, *supra* note 10, at 218 (concluding that "only the Supreme Court had anything good to say about scarcity" in the 1970s (quoting Daniel Polsby, *Candidate Access to the Air*, 1981 SUP. CT. REV. 223)). In *League of Women Voters*, the Court acknowledged the prevalent criticism of the scarcity doctrine and stated that reevaluation would require "some signal from Congress or the FCC that technological developments have advanced so far that some revision of the system of broadcast regulation may be required." FCC v. League of Women Voters of Cal., 468 U.S. 364, 376 n.11 (1984). The Court reiterated this position in *Turner I*. Fox Television Stations, Inc. v. FCC, 280 F.3d 1027, 1046 (D.C. Cir. 2002) (stating that "[t]he Supreme Court has already heard the empirical case against" the scarcity doctrine, and still "declined to question its continuing validity" (quoting Turner Broad. Sys., Inc. v. FCC, 512 U.S. 622, 638 (1994))).

Professor Hazlett has suggested that the scarcity doctrine's inherent ambiguity is itself the reason for its durability. Hazlett, *supra* note 77, at 929. By characterizing scarcity as an "objective fact," without addressing competing technologies, the *Red Lion* opinion makes "empirical falsification" impossible. *Id.*

96. *See* FCC v. Nat'l Citizens Comm. for Broad., 436 U.S. 775, 795 (1978).

97. *See League of Women Voters*, 468 U.S. at 374-75.

three decisions from the 1990s summarize the Court's current position on broadcast scarcity.

In *Metro Broadcasting, Inc. v. FCC*, the Court considered FCC regulations enhancing licensing opportunities for minority owners.[98] In upholding the regulations, the Court acknowledged that scarcity justifies governmental restraints on licensees that favor both viewpoints and speakers.[99] The Court viewed the need for government selection of broadcasters as "axiomatic," citing and echoing the scarcity arguments made in *NBC* and *Red Lion*.[100] Nearly fifty years after its birth, the scarcity doctrine received strong affirmation in *Metro Broadcasting*.

Limits to the scarcity theory, however, emerged in two later decisions. In *Turner Broadcasting System, Inc. v. FCC* (*Turner I*), the Court analyzed a statute requiring cable television providers to transmit local broadcast television channels to subscribers without charge.[101] The Court, in dicta, stated that the "less rigorous" First Amendment scrutiny applied to broadcast regulations is premised on "the unique physical limitations of the broadcast medium."[102] The Court then distinguished electromagnetic broadcasting from cable television stating, "cable television does not suffer from the inherent limitations that characterize the broadcast medium."[103] Any possible physical limitations in cable broadcasting, therefore, are insufficient to alter the normal protections of the First Amendment.[104]

Three years later, in *Reno v. ACLU*, the Court used the reasoning of *Turner I* to distinguish the Internet from broadcast television.[105] The Court found that the Internet could not be considered a scarce resource, given its ability to provide a "relatively unlimited, low-cost capacity for communication of all kinds."[106] The Internet, the Court illustrated, can transform any speaker into a "town crier" and any user into "a pamphleteer."[107] Relaxed First Amendment scrutiny was therefore unnecessary.[108]

Conclusions about the current state of the scarcity doctrine are difficult. The decisions in *Metro Broadcasting* and *Turner I* affirming the vitality of the doctrine have aged rapidly during the explosive growth of new media at the close of the last century. Moreover, the Court's piecemeal exclusion of cable and Internet broadcasting appears to be on a collision course with science. Finally,

---

98. 497 U.S. 547, 552 (1990).

99. *Id.* at 566-67.

100. *Id.* at 567.

101. Turner Broad. Sys., Inc. v. FCC, 512 U.S. 622, 631-32 (1994).

102. *Id.* at 637. The Court noted, however, that the broadcast cases "are inapposite" to the cable television context. *Id.* at 638-39.

103. *Id.* at 639.

104. *Id.*

105. 521 U.S. 844, 869-70 (1997).

106. *Id.* at 870.

107. *Id.*

108. *Id.*

the acknowledgment in *Turner I* that convergence has removed old distinctions within the various broadcast media[109] would appear to undermine whatever remaining analytic force scarcity once held. Taken together, these decisions suggest that the physical scarcity doctrine will not satisfy future regulations on broadcast speech.[110]

D. Fairness and Equal Time: The Regulation of Broadcast Politics

Scarcity theories are important to a discussion of the broadcasting of presidential debates because the regulation of broadcast media has developed through interaction between the FCC and the Supreme Court.[111] The deferential scrutiny applied to the FCC's broadcast policies has permitted a wide range of regulations addressing social issues such as indecency[112] and diversity.[113] Likewise, the FCC has promulgated a series of regulations designed to increase public involvement in the democratic process and political elections.[114]

Section 315(b)(1) of the Communications Act of 1972, for instance, allows candidates for political office to purchase broadcast airtime at the "lowest unit charge" offered to other purchasers for the same time and period.[115] The lowest unit charge rule was intended to prevent broadcasters from exercising their market power to extract additional profits from candidates and to maintain the availability of the broadcast forum.[116]

A second doctrine required broadcasters to give candidates for federal office

---

109. *Id.* at 627.

110. Then Commissioner, now Chairman of the FCC, Michael K. Powell voiced similar doubts on the modern (as well as historical) values of the scarcity rationale, commenting "that if scarcity was ever a defensible explanation it is certainly farcical in the modern digital era, which is marked by abundance." Michael K. Powell, Remarks Before the Media Institute, Accepting Freedom of Speech Award (Oct. 20, 1999), *available at* http://www.fcc.gov/Speeches/Powell/ spmkp905.html (last visited Mar. 17, 2002).

111. *See* BOLLINGER, *supra* note 9, at 66.

112. *See generally* FCC v. Pacifica Found., 438 U.S. 726 (1978).

113. FEDERAL COMMUNICATIONS COMMISSION, PUBLIC SERVICE RESPONSIBILITY OF BROADCAST LICENSEES 15 (1946) ("[I]t has long been an established policy of . . . the Commission that the American system of broadcasting must serve significant minorities among our population, and the less dominant needs and tastes . . . .").

114. STEVEN J. SIMMONS, THE FAIRNESS DOCTRINE AND THE MEDIA 72-80 (1978).

115. 47 U.S.C. § 315(b)(1) (1994).

116. *See* Angela J. Campbell, *Political Campaigning in the Information Age: A Proposal for Protecting Political Candidates' Use of On-Line Computer Services*, 38 VILL. L. REV. 517, 550 (1993). Former Chairman of the FCC Reed Hundt has argued that the rule "fails in practice" because broadcasters steer candidates to higher-priced time periods. Press Release, FCC, FCC Chairman Reed Hundt Calls on FCC to Launch Major Free Time Initiative (Sept. 12, 1997), *available at* http://www.fcc.gov/Bureaus/Miscellaneous/News_Releases/1997/nrmc7065.html (advocating the use of "low, even zero" cost rates for candidates with a cap on the amount of airtime that could be purchased).

"reasonable access" to broadcast time to advocate their candidacy.[117] The reasonable access doctrine sought to increase voter education through broadcast appearances[118] and alleviate concerns over insufficient broadcast coverage.[119] Despite the intrusion into the broadcasters' editorial decisions, the Supreme Court upheld the reasonable access rules as a permissible licensing condition.[120]

In addition, the FCC has long required broadcasters to provide equal access to the airwaves to all "legally qualified"[121] political candidates, when any one of them is granted broadcast time. The equal time provision was first adopted in the Radio Act of 1927[122] and carried over into the 1934 Act[123] to prevent the potential bias of a broadcast station providing exclusive coverage to a single candidate.[124] In 1959, Congress amended the statute to exempt news coverage of candidates,[125] swiftly rejecting the FCC's narrower interpretation of the statute.[126] In 1960, Congress temporarily suspended the equal time rule[127] to permit the first televised

---

117. 47 U.S.C. § 312(a)(7) (1990).

118. S. Rep. No. 92-96, at 20 (1971), *reprinted in* 1972 U.S.C.C.A.N. 1773, 1774 (noting that Congress intended "to give candidates for public office *greater access to the media* so that they may better explain their stand on the issues, and thereby more fully and completely inform the voters") (emphasis in original). *But see infra* notes 175-76 and accompanying text.

119. *See* Kennedy for President Comm. v. FCC, 636 F.2d 432, 442 (D.C. Cir. 1980).

120. CBS v. FCC, 453 U.S. 367 (1981).

121. 47 U.S.C. § 315(a) (1994). The 1934 Act does not define which candidates are "legally qualified," and thus the applicability of § 315(a) depends on state, federal, or local law requirements for candidacy. FCC, Rules Applicable to All Broadcast Stations, 47 C.F.R. § 73.1940 (2001); *see also* Lili Levi, *Professionalism, Oversight, and Institution-Balancing: The Supreme Court's "Second Best" Plan for Political Debate on Television*, 18 YALE J. ON REG. 315, 376 n.193 (2001).

122. Radio Act of 1927, Pub. L. No. 69-632, 44 Stat. 1162, 1170.

123. 47 U.S.C. § 315 (1982).

124. Thomas Blaisdell Smith, Note, *Reexamining the Reasonable Access and Equal Time Provisions of the Federal Communications Act: Can These Provisions Stand if the Fairness Doctrine Falls?*, 74 GEO. L.J. 1491, 1497 (1986).

125. The amendment was prompted by the FCC's decision that network coverage of routine news events involving one candidate for office triggered equal time obligations for all other qualified candidates. SIMMONS, *supra* note 114, at 46-47 (discussing the FCC's decision in *In re* Petitions of CBS and NBC for Reconsideration and Motions for Declaratory Rulings or Orders Relating to the Applicability of § 315 of the Communications Act of 1934, as amended, to Newscasts by Broadcast Licensees, Interpretive Op., 26 F.C.C. 715 (1959)).

126. KRATTENMAKER & POWE, *supra* note 10, at 67.

127. Act of Aug. 24, 1960, Pub. L. No. 86-677, 74 Stat. 554. Broadcasters lobbied for the suspension of the equal time regulation in part because the amendments to the equal access doctrine in 1959 failed to clarify broadcasters' duties under the statute. *See* ERIK BARNOUW, THE IMAGE EMPIRE: A HISTORY OF BROADCASTING IN THE UNITED STATES VOL. III 161-62 (1970). The lobbying effort was also launched because of lingering network hostility to the duties imposed by the equal time rule. In particular, the networks were "unwilling to give time away to the major parties' presidential candidates under circumstances that would force them to give equal opportunities to numerous minor-party candidates." David M. Rice, *Network Television as a*

presidential debates.[128] Finally, in 1975, the Commission dismantled the equal time rule by classifying political campaign debates as "bona fide news events" within the 1934 Act's exception,[129] paving the way for modern television election coverage.[130]

However, the fairness doctrine, the most sweeping restriction designed by the FCC, still impeded full media coverage of election politics. The fairness doctrine, in the words of the Supreme Court, "imposed on radio and television broadcasters the requirement that discussion of public issues be presented on broadcast stations, and that each side of those issues must be given fair coverage."[131] Less a doctrine than a direction, the fairness concept grew from the earliest days of the FRC through a series of individual complaints and rulings.[132] Congress formally, but perhaps unintentionally, adopted the doctrine in the 1957 amendments to the equal time provisions.[133] As codified, the fairness doctrine required broadcasters to air all sides of controversial public issues, irrespective of the broadcasters' own interest in covering such material.[134] Again citing the scarcity theory, the Supreme Court upheld the fairness doctrine as constitutional and extended the government unique latitude to regulate broadcasting.[135] Despite swift and sustained criticism,[136] the fairness doctrine remained in force until the

---

Medium of Communication, *in* NETWORK TELEVISION AND THE PUBLIC INTEREST 198 (Michael Botein & David M. Rice eds., 1980).

128. *See infra* notes 160-71 and accompanying text.

129. 47 U.S.C. § 315(a)(4) (1994). The bona fide news event rule exempted four types of political broadcasts from the reasonable use rule, including any "bona fide newscast," "bona fide news interview," "bona fide news documentary," or "on the spot coverage of bona fide news events" such as political conventions. *Id.* § 315(a)(1)-(4). Following the 1960 debates, the Commission ruled that political debates sponsored by "nonbroadcast entities" or independent third parties, covered live, were bona fide news events within the statutory exemption. *In re* Petitions of the Aspen Inst. Program on Communications and Soc'y & CBS, Inc. for Revision or Clarification of Comm'n Rulings Under Section 315(a)(2) & 315(a)(4), 55 F.C.C.2d 697 (1975), *aff'd sub nom.*, Chisholm v. FCC, 538 F.2d 349 (D.C. Cir. 1976). The "nonbroadcast entity" requirement was abandoned in 1983. *In re* Petitions of Henry Geller & Nat'l Assoc. of Broads. & the Radio-Television News Dirs. Assoc. to Change Comm'n Interpretation of Subsections 315(a)(3) and (4) of the Communications Act, 95 F.C.C.2d 1236 (1983), *aff'd sub nom.*, League of Women Voters Educ. Fund v. FCC, 731 F.2d 995 (D.C. Cir. 1984).

130. *See* Youm, *supra* note 4, at 695.

131. Red Lion Broad. Co. v. FCC, 395 U.S. 367, 369 (1969). The extensive scholarly attention devoted to the fairness doctrine is outside the scope of this Article.

132. Thomas G. Krattenmaker & L.A. Powe, Jr., *The Fairness Doctrine Today: A Constitutional Curiosity and an Impossible Dream*, 1985 DUKE L.J. 151, 152 n.7.

133. *Red Lion*, 395 U.S. at 380 (concluding "the amendment vindicated the FCC's general view that the fairness doctrine inhered in the public interest standard" of § 315).

134. BRUCE M. OWEN, ECONOMICS AND FREEDOM OF EXPRESSION: MEDIA STRUCTURE AND THE FIRST AMENDMENT 116 (1975).

135. *Red Lion*, 395 U.S. at 375.

136. Not surprisingly, much of the criticism was generated by the broadcast industry itself.

mid-1980s, when the FCC agreed to revisit the analytical foundation for the rule.

In 1985, the FCC responded to criticism of the fairness doctrine, issuing a report on its continuing viability.[137] The Commission found that the fairness doctrine worked to limit broadcast coverage of controversial issues in order to minimize the amount of reply time devoted to opposing sides of a public concern.[138] The FCC further concluded that the theory of spectrum scarcity that supported the *Red Lion* decision no longer justified "per se" regulation of broadcasters, "particularly rules which affect the constitutionally sensitive area of content . . . ."[139] The FCC conceded that the fairness doctrine was an "unnecessary and detrimental regulatory mechanism," given the growth of new information sources, the intrusion into broadcast editorial privileges, and the lack of a demonstrated public benefit.[140] The FCC soon formally abandoned the fairness doctrine,[141] with the last two small public interest duties, the personal attack[142] and political editorial rules,[143] repealed by writ of mandamus in 2000.[144]

The philosophical bases for broadcast regulations thus reveal both historical misconceptions and modern inconsistencies. The scarcity rationale, questioned as scientifically flawed from its inception,[145] is now clearly minimized by the

---

See BEN H. BAGDIKIAN, THE MEDIA MONOPOLY 247 (6th ed. 2000).

137. General Fairness Doctrine Obligations of Broadcast Licensees, 50 Fed. Reg. 35,418 (Aug. 30, 1985).

138. *Id.* at 35,423. The Commission further noted that broadcasters faced heavy economic burdens in complying with the fairness doctrine including costs of defending against administrative challenges. *Id.* at 35,435. While the costs alone were not sufficient to justify eliminating the doctrine, no "counterveiling justifications" offset the hardships, as no additional public interest programming was produced. *Id.*

139. *Id.* at 35,422. The Commission described the constitutional concerns as contravening "fundamental constitutional principles" and according a "dangerous opportunity for governmental abuse . . . ." *Id.* at 35,446.

140. *Id.* at 35,445-46.

141. Meredith Corp. v. FCC, 809 F.2d 863 (D.C. Cir. 1987), *on remand*, *In re* Compl. of Syracuse Peace Council Against Television Station WTVH Syracuse, New York, 2 F.C.C.R. 5043 (1987).

142. The personal attack rule stated that "if an attack is made on someone's integrity during a presentation of views on a controversial issue of public importance, the licensee must inform that person . . . and provide a reasonable opportunity to respond." *In re* Repeal or Modification of the Personal Attack and Political Editorial Rules, 15 F.C.C.R. 19,973, 19,973 n.2 (2000) (citing 47 C.F.R § 73.1920 (2000)).

143. The political editorial rule stated that if a broadcast station airs an editorial supporting a "legally qualified candidate," the broadcaster must provide a "reasonable opportunity" for opposing candidates to respond. *Id.* (citing 47 C.F.R. § 73.1930 (2000)).

144. Radio-Television News Dirs. Ass'n v. FCC, 229 F.3d 269 (D.C. Cir. 2000). The nearly two-decade struggle over the fate of the fairness doctrine is beyond the scope of this Article. A helpful summary of the battle is presented in Ian Heath Gershengorn, *The Fall of the FCC's Personal Attack and Political Editorial Rules*, 19 COMM. LAW 7 (2001).

145. Hazlett, *supra* note 77, at 926-31.

development of broadband, cable television, satellite, the Internet, and proprietary on-line networks.[146] As competition grows in the media marketplace, consumers will find substitutes to the network monopolies that satisfy their once ignored niche tastes. Competition, therefore, will produce the diverse array of speakers promised by regulation but never achieved.

In the narrow context of political campaigns, however, the public interest principle retains more viability. The growth of broadcast network alternatives has stimulated content competition, and the entry of new competitors has decreased the marginal costs of programming. But broadcast technology remains costly,[147] and new consolidation within the broadcast markets threatens to restrict market entry anew.[148] One premise of the 1927 Act thus remains relevant, as unequal access to broadcast technology could unfairly advantage a single candidate.[149] Particularly in the general presidential elections, wide access to

---

146. *See supra* notes 77-89 and accompanying text.

147. *See* Debora L. Osgood, Note, *Expanding the Scarcity Rationale: The Constitutionality of Public Access Requirements in Cable Franchise Agreements*, 20 U. MICH. J.L. REFORM 305, 327-32 (1986) (advocating the application of the scarcity doctrine to cable television by analyzing scarcity in an economic, rather than physical context).

148. *See, e.g.*, Time Warner Entm't Co. v. FCC, 240 F.3d 1126 (D.C. Cir. 2001). In *Time Warner*, the D.C. Circuit found that FCC rules capping the number of subscribers serviced by a cable television company and limiting the amount of programming produced by the cable company that may be shown on its own networks violate the cable companies' First Amendment rights. *Id.* The decision allows major cable providers such as AOL Time Warner to expand their national markets, possibly at the expense of independent producers. Stephen Labaton & Geraldine Fabrikant, *U.S. Court Ruling Lets Cable Giants Widen Their Reach*, N.Y. TIMES, Mar. 3, 2001, at A1. Similarly, the FCC modified broadcast regulations to allow the four major broadcast networks to own or operate emerging networks such as UPN or the WB. Press Release, FCC, FCC Eliminates the Major Network/Emerging Network Merger Prohibition from Dual Network Rule (Apr. 19, 2001) (on file with author).

149. Commentators have noted that the Supreme Court's decision in *NBC*, and the development of the scarcity doctrine itself, may have been motivated by this concern. P.M. Schenkkan, Comment, *Power in the Marketplace of Ideas: The Fairness Doctrine and the First Amendment*, 52 TEX. L. REV. 727, 742 (1974). Schenkkan argues that the Court used scarcity as a means of explaining its true concern, a monopolization of the broadcast medium by a few, favored speakers. *Id.* In support, he notes Justice Murphy's dissent in *NBC* centered not on spectrum scarcity alone, but on the danger of allowing the scarce spectrum to become "a weapon of authority and misrepresentation." *Id.* at 742-43 (quoting NBC v. United States, 319 U.S. 190, 228 (1943) (Murphy, J., dissenting)).

Scholars, however, have also noted that any perceived concentration of power in the hands of a few network broadcasters "has been broken by deregulation and technology." KRATTENMAKER & POWE, *supra* note 10, at 222. Moreover, concentrating broadcast regulation in the hands of a single administrative monopoly is not necessarily a more satisfying protection against possible abuse. *See* HUBER, *supra* note 90, at xiv ("[I]n 1934 . . . the United States folded all federal authority over both wireline and wireless communication into a new, superpowerful communications commission . . . . Germany got an FCC too, even bigger and more effective than

information on the candidates remains essential to our system of participatory, indirect democracy.

II. DEBATES IN THE MODERN PRESIDENTIAL ELECTION

Within the hierarchy of speech values,[150] the Supreme Court has singled out campaigns for public office as a core value protected by the First Amendment.[151] Candidates for public office engage in a wide variety of direct and indirect speech, including rallies, fundraisers, and orchestrated news events.[152] Elevated above all these events, however, the candidate debates occupy a central place in American politics.[153]

A. Presidential Debates, Before and After Television

Candidate debates date back to at least 1788, when James Madison campaigned for election to the House of Representatives.[154] History records the epic confrontations between Stephen A. Douglas and Abraham Lincoln, when a captivated nation listened to the candidates duel on the future of slavery, unification, and federal governance.[155] As answers to contemporary social questions developed into political party allegiance, the public debates provided a peaceful forum for airing disputes within government.[156]

---

ours. . . . Our own commission never got that bad, but it got bad enough . . . ."). Huber's allusion, however melodramatic, does point to the problem: "[H]istory teaches that the fear said to justify regulation of speech exists all too often only in the minds of the regulators." KRATTENMAKER & POWE, *supra* note 10, at 224.

150. Although the Court has never formally established a hierarchy of First Amendment values, it has acknowledged a tiered value system on numerous occasions. *E.g.*, Connick v. Myers, 461 U.S. 138, 145 (1983) (noting that "speech on public issues occupies the 'highest rung of the hierarchy of First Amendment values, and is entitled to special protection" (quoting Carey v. Brown, 447 U.S. 455, 467 (1980))).

151. Monitor Patriot Co. v. Roy, 401 U.S. 265, 271-72 (1971) ("[I]t can hardly be doubted that the constitutional guarantee has its fullest and most urgent application precisely to the conduct of campaigns for political office."). The Court has reiterated this position in *McIntyre v. Ohio Elections Commission*, 514 U.S. 334, 347 (1995), and *Buckley v. Valeo*, 424 U.S. 1, 14-15 (1976).

152. Mark C. Alexander, *Don't Blame the Butterfly Ballot: Voter Confusion in Presidential Politics*, 13 STAN. L. & POL'Y REV. 121 (2002).

153. The Supreme Court has stated that "[d]eliberation on the positions and qualifications of candidates is integral to our system of government, and electoral speech may have its most profound and widespread impact when it is disseminated through televised debates." Ark. Educ. Television Comm'n v. Forbes, 523 U.S. 666, 676 (1998); *see also* Raskin, *supra* note 4, at 1944.

154. KATHLEEN HALL JAMIESON & DAVID S. BIRDSELL, PRESIDENTIAL DEBATES: THE CHALLENGE OF CREATING AN INFORMED ELECTORATE 34 (1988).

155. *See generally* THE LINCOLN-DOUGLAS DEBATES (Harold Holzer ed., 1993).

156. JAMIESON & BIRDSELL, *supra* note 154, at 40 ("[D]ebate[s] became a release valve for the pressures of constituency and faction . . . . [T]he cycle of debate offered clash without disaster and simultaneously affirmed the value and legitimacy of the political structure that made debate

Although political debates had long been printed and distributed along the campaign trail,[157] broadcasting promised a new era of open democracy. Herbert Hoover's insistence on radio regulations designed for public benefit drew upon his belief that broadcasting would revolutionize political debates.[158] Until the emergence of broadcasting, the candidates often viewed presidential debates as dangerous. Broadcasting changed this pattern, allowing candidates to speak to a national audience.[159]

The arrival of television transformed the presidential debates into the seminal event in election politics.[160] Television seemed to hold the power to electrify American politics, reuniting citizens with Washington by providing live access to the candidates.[161] Although televised political coverage was common by 1960, presidential candidates were still wary of violating the age-old maxim against appearing alongside a rival.[162] Broadcasters, however, saw an important public service opportunity in providing free airtime to the candidates.[163] In 1960, eager

---

possible.").

157. *Id.* at 51 (explaining that candidates such as Lincoln and Daniel Webster published debate speeches in pamphlets and newspaper editorials).

158. *Id.* at 84 ("Hoover believed that radio would revolutionize 'the political debates that underlie political action [by making] us literally one people upon all occasions of general public interest.'"). *See also* Krasnow & Goodman, *supra* note 44, at 608 (quoting Hoover's belief that the "ether is a public medium, and its use must be for public benefit").

159. In the early years of radio, presidential candidates largely ignored the possibilities offered by the new broadcast format. *See* CNN, THE DEBATES '96, PRESIDENTIAL DEBATE HISTORY, HOW WE GOT THEM, AND WHAT THEY MEAN, *at* http://cgi.cnn.com/ALLPOLITICS/ 1996/debates/ history [hereinafter CNN, THE DEBATES '96] (last visited Apr. 19, 2002). Prior to the debates of 1960, only one presidential election debate was broadcast over radio, a single Republican primary contest in Oregon between Thomas Dewey and Harold Stassen. *Id.*; *see also* COMMISSION ON PRESIDENTIAL DEBATES, DEBATE HISTORY: 1948 DEBATES, *at* http://www.debates. org/pages/ debhis48.html (last visited Apr. 12, 2002) (estimating forty to eighty million listeners tuned in to the one hour debate on outlawing the Communist Party in the United States).

160. Keith Darren Eisner, Comment, *Non-Major Party Candidates and Televised Presidential Debates: The Merits of Legislative Inclusion*, 141 U. PA. L. REV. 973, 974-75 (1993). The first televised presidential debate occurred during the 1956 Democratic primary in Florida between candidates Adlai Stevenson and Estes Kefauver. COMMISSION ON PRESIDENTIAL DEBATES, DEBATE HISTORY: 1956 DEBATES, *at* http://www.debates.org/pages/debhis56.html (last visited Apr. 12, 2002). For a detailed account of the debate, see JAMIESON & BIRDSELL, *supra* note 154, at 92-93.

161. Angus Campbell, *Has Television Reshaped Politics?*, *at* http://www.mbcnet.org/ debateweb/html (last visited Apr. 5, 2002) (quoting former CBS president Dr. Frank Stanton as remarking "[t]elevision, with its penetration, its wide geographic distribution and impact, provides a new, direct and sensitive link between Washington and the people").

162. Earl Mazzo, *The Great Debates*, *at* http://www.mbcnet.org/debateweb/html (last visited Apr. 6, 2002); *see also* CNN, THE DEBATES '96, *supra* note 159 (noting that prior to 1960 the "most vocal group advocating debates were political underdogs wanting to share the stage with incumbents").

163. CNN, THE DEBATES '96, *supra* note 159.

to convince Congress to eliminate the equal-time provision of the 1934 Act,[164] the major networks volunteered dozens of free hours to each candidate in the weeks preceding the general election.[165] After private negotiations, the networks agreed to a live unified broadcast without sponsorship or commercial interruption.[166] United by the possible permanent repeal of the equal-time rule,[167] the level of cooperation among the networks to air the unprecedented event was unusually high. As a result, when Kennedy and Nixon took to the stage, nearly every television station in the country carried the event.[168]

The national reaction to the first Kennedy-Nixon debate was enormous[169]—a record audience of over sixty-six million households.[170] Richard Nixon proclaimed that "debates between the presidential candidates are a fixture," and predicted that "in all the elections in the future we are going to have debates."[171] Yet televised presidential debates nearly disappeared from the political landscape in subsequent years.[172] No presidential debates were held between 1960 and 1976,[173] and the televised debate re-emerged only after painstaking negotiations between the candidates.[174] Equally disappointing was the impact of the televised

---

164. *See supra* notes 122-30 and accompanying text.

165. Editorial, *Senate Suspends § 315, Now it's up to the House to Follow Suit*, BROADCASTING, July 4, 1960, *available at* http://www.mbcnet.org/debateweb/html/ history/1960/ section315.htm [hereinafter Editorial, *Senate Suspends*] (last visited Apr. 2, 2002).

166. Editorial, *Sponsorship of TV Debates?*, BROADCASTING, Aug. 8, 1960, *available at* http://www.mbcnet.org/debateweb/html/history/1960/sponsorship.htm (last visited Apr. 2, 2002). Although prominent advertisers lined up to provide exclusive sponsorship, CBS seized the public relations opportunity to decline the revenue, proclaiming "we . . . want to make this our own contribution because we believe there is no single act of self-government that is more important than the quadrennial choice of our national leadership." *Id.* The other networks quickly followed.

167. *See* Editorial, *Senate Suspends*, *supra* note 165 (stating that the suspension of the equal-time provision required the FCC to report on the results of the suspension during the 1960 election and to "recommend any legislation it thinks necessary" to repeal the rule permanently).

168. Mazzo, *supra* note 162 ("Almost every station carried the debates simultaneously, and in most places there were no alternative programs.").

169. Special Report, *"Great Debate" Rightly Named, Nixon, Kennedy Set a Precedent That Will Be Hard to Abandon*, BROADCASTING, Oct. 3, 1960, at 88, *available at* http://www.mbcnet.org/ debateweb/html/history/1960/rightlynamed.htm (last visited Apr. 5, 2002) (proclaiming the "whole course of political campaigning has been changed by a single broadcast").

170. COMMISSION ON PRESIDENTIAL DEBATES, DEBATE HISTORY: 1960 DEBATES, *at* http://www.debates.org/pages/debhis60.html (last visited Mar. 20, 2002).

171. STEPHEN BATES, THE FUTURE OF PRESIDENTIAL DEBATES (1993), *at* http://www.annenberg.nwu.edu/pubs/debate (last visited Apr. 1, 2002). *But see infra* note 179.

172. Susan E. Spotts, *The Presidential Debates Act of 1992*, 29 HARV. J. ON LEGIS. 561, 563 (1992).

173. CNN, THE DEBATES '96, *supra* note 159.

174. BATES, *supra* note 171 (noting that the 1976 debates between Jimmy Carter and Gerald Ford required "haggling over timing, format, questioners, camera angles, risers, notes, stools, props, and a host of other issues").

debates on political involvement. The televised debates, as well as television coverage of politics and campaigns in general, did not increase voter turnout[175] or voter interest in the elections.[176] The much-anticipated revolution in American democracy, it appeared, would not be televised. The seeming failure of televised debates to invigorate the electorate defies a single explanation. Instead, the problems of the televised debates are political, technological, and historical.

First, presidential candidates may have no political incentive to debate their opponents. The great debates of 1960 occurred only because the broadcast networks lobbied the candidates, Congress, and the FCC intensively. Vice President Nixon enjoyed the advantages of national recognition and understood that Senator Kennedy would benefit from merely sharing the stage.[177] Then, as now, a presidential underdog could win valuable momentum by merely "holding his own" in the debates and avoiding "visibly serious blunders."[178] In many contests, therefore, the leading candidate has more to lose in the debates and a strong motivation to decline a televised confrontation.[179]

Second, candidates tend to narrow their messages during televised debates

---

175. Campbell, *supra* note 161. Campbell reported on research concerning voter turnout in the era of televised politics. Noting that the "most commonly accepted indicator of public involvement in politics is the turnout in national elections," Campbell found that:

> In fact, there has been only a slight rise in the turnout figures during the last ten years. In the presidential elections of 1952, 1956, and 1960 the turnouts—that is, the proportions of adult citizens who voted—were considerably higher than in the elections of 1944 and 1948, but if we drop back to the period just before the war we find that the turnouts in 1936 and 1940 were almost as high as they have been in the most recent elections. There has been a small proportionate increase in the presidential vote during the television era, although it has fluctuated and at its lowest point in 1956 (60.4 percent) exceeded by only a percentage point the high of the pre-television period.

Id.

176. *Id.* Campbell presented the findings of an ongoing study investigating voter interest between 1952 and 1960. The sampling found a large fluctuation of voter interest, but also noted a "tremendous increase in television coverage" during these same years. The findings, Campbell explains, are important because if "television had demonstrated a unique capacity to activate political interest among its viewers we should find a substantial increase in the number expressing high interest over the 1952 to 1960 period. This we do not find." *Id.*

177. Mazzo, *supra* note 162 (noting then Vice President Nixon was acutely aware of his advantage over the relatively unknown Senator Kennedy, and thus reluctant to provide prestige to a lesser-known opponent).

178. *Id.* (discussing Senator Kennedy's advantage in the 1960 debates). Forty years later, Republican candidate George W. Bush enjoyed the same advantage against his more seasoned rival Vice President Albert Gore, Jr. Richard L. Berke, *Debates Put in Focus Images and Reality*, N.Y. TIMES, Oct. 19, 2000, at A29 (quoting democratic strategist David Axelrod who noted "I think Bush gained the most of [the debates] just by surviving").

179. BATES, *supra* note 171. Thus, despite his enthusiasm for televised debates in 1962, Richard Nixon refused to debate his challengers in 1968 and 1972, concluding "[i]t's poor tactics when you're running so far ahead." *Id.* (quoting Spiro Agnew).

to appeal to a national audience. Traditionally, non-televised debates have been "free-flowing," without pre-set questions and intrusive moderation.[180] This unstructured, adversarial format was a product of the limited audience. Like most political rallies, a close-knit network of insiders usually attended the live debates.[181] Broadcasting, in contrast, cuts across the social spectrum reaching all levels of income, education, and political involvement.[182] Candidates gained access to a national audience, but could no longer assume they spoke only to the faithful. The candidates responded by modifying their message for the broadcast medium. The adversarial format of the traditional debate was merged with question-and-answer sessions[183] and press conferences[184] to create an often stale mix of substance and showmanship. Not surprisingly, viewers frequently found programming alternatives more appealing.[185]

Third, the televised debates are burdened by the inherent limits of television. Television viewing is a largely passive activity.[186] Its importance in the political process is often "the ease with which television news falls into its audience's laps. . . ."[187] Candidates are forced into the difficult position of finding the highest plane of dialogue consistent with the education and interest of the audience.[188] Unable to easily define this target, candidates concluded that winning the televised debates requires satisfying the media instead of the viewers.[189] In turn, television journalism, predisposed to "drama and visual

---

180. JAMIESON & BIRDSELL, *supra* note 154, at 87.

181. *See* PAUL TAYLOR, SEE HOW THEY RUN, ELECTING THE PRESIDENT IN AN AGE OF MEDIAOCRACY 245 (1990); *see also* Alexander, *supra* note 152, at 125 (noting campaign rallies are "packed with supporters" and "designed to motivate the faithful").

182. TAYLOR, *supra* note 181, at 244.

183. JAMIESON & BIRDSELL, *supra* note 154, at 102.

184. *Id.* at 118.

185. The problem of viewer attrition is not new. Researchers noted that even viewership of the great debates of 1960 waned in areas where alternatives were broadcast. *See* Mazzo, *supra* note 162 (estimating viewership of the debates dropped between fifteen and twenty percent in areas where local affiliates carried alternatives to the first Nixon/Kennedy debate).

186. TAYLOR, *supra* note 181, at 244. Taylor notes that "[w]atching television is a passive, low-intensity activity—'chewing gum for the eyes'—which requires less concentration than reading a book or newspaper." *Id.*

187. *Id.*

188. JAMIESON & BIRDSELL, *supra* note 154, at 15 ("The audience for presidential debating is far less directed and accountable . . . . The audience probably employs some set of standards, but these are informal and inexplicit. . . ."). Taylor thus concludes that the television audience "is broader, less educated, less sophisticated and less interested in public affairs than the readership of newspapers." TAYLOR, *supra* note 181, at 244.

189. *See* Bob Davis & Jackie Calmes, *Debaters Decoded: A Viewer's Guide to Tomorrow's Words*, WALL ST. J., Oct. 2, 2000, at A1. The authors summarize the goals of the modern presidential candidate in each televised debate: "to introduce themselves to Americans who have been too bored to pay attention to the presidential race, avoid embarrassing missteps, and make each other look bad." *Id.*

imagery,"[190] tends to focus on which candidate "won" the contest rather than the substance of the issues discussed.[191] Candidates fearing the stigma that accompanies a perceived "loss" are forced to spend countless hours preparing for each debate in an effort to appear "poised and confident."[192] The viewers are then treated to endless predictions and post-debate opinion polls,[193] little of which assists an informed debate on the candidates' ability to govern.[194]

Finally, televised debates often offer nothing new. Commentators have noted that the "essential problem of all political communication is the character of the public demand for it."[195] Television, like all media, has the capacity to reach a demographically diverse audience and thus augment the education of all voters.[196] Researchers observe, however, that the primary consumers of political television, including the presidential debates, are usually the most informed segments of society.[197] Social scientists agree, noting that televised debates largely reinforce voter preference.[198] The debate audience, therefore, is frequently comprised of the same group that follows the election most closely in other media such as newspapers and radio.[199]

---

190. LYNDA LEE KAID, POLITICAL PROCESS & TELEVISION, MUSEUM OF BROADCASTING AND COMMUNICATIONS, ENCYCLOPEDIA OF TELEVISION, *available at* http://www.mbcnet.org/debateweb/html (last visited Apr. 5, 2002).

191. *See, e.g.*, Editorial, *Winner of Debate is American Public*, CHI. SUN-TIMES, Oct. 4, 2000, at 55 (concluding the first 2000 presidential debate failed to "settle the presidential election" because neither "candidate scored a knockout, and neither committed a candidacy-killing mistake").

192. PBS, DEBATING OUR DESTINY: PREPARING FOR THE DEBATES, *available at* http://www.pbs.org/newshour/debatingourdestiny/debate-prepping.htm (last visited Apr. 10, 2002) (on file with author); *see also* Alexander, *supra* note 152, at 127 (noting that preparing for the debates "is one of the most intense exercises that a campaign endures").

193. *See* Jackie Calmes & Jeanne Cummings, *Bush Tries to Score a Few Points After the Bell; Polls Give Gore Debate Edge, but Rival Pounces on Exaggerations*, WALL ST. J., Oct. 5, 2000, at A28 (discussing the results of "[s]nap opinion polls" favoring the Vice-President's performance in the first televised debate).

194. *See* Richard L. Berke & Kevin Sack, *In Debate 2, Microscope Focuses on Gore*, N.Y. TIMES, Oct. 11, 2000, at A1.

195. Campbell, *supra* note 161.

196. *Id.*

197. *Id.*

198. *See* Peter R. Schorott, *Electoral Consequences of "Winning" Televised Campaign Debates*, 54 PUB. OP. Q. 567, 568 (1990) (citing research indicating that "voters adopted the issue position taken by their preferred candidate" following broadcast debates). Social scientists have labeled this process as "'group polarization' in which like-minded people in an isolated group reinforce one another's views, which then harden into more extreme positions." Alexander Stille, *Adding Up the Costs of Cyberdemocracy*, N.Y. TIMES, June 2, 2001, at B9. Televised debates help overcome these "selective attention barriers" by exposing the partisan audience to the views of the opposing candidate. JAMES B. LEMERT ET AL., NEWS VERDICTS, THE DEBATES, AND PRESIDENTIAL CAMPAIGNS 199 (1991).

199. Campbell, *supra* note 161. Campbell cites this overlap in viewership to explain the small

Each of these problems has contributed to a steady decline in presidential debate viewership.[200] By 1992, the presidential debates were no longer "must see television" as viewership dropped to under thirty-seven million households, fewer than half the homes of just sixteen years earlier.[201] While the presidential debates still serve an important function,[202] they have not revolutionized the substance or structure of the presidential election.[203]

B. The 2000 General Election Debates

During the 2000 elections, the networks finally lost interest. Faced with declining ratings and vigorous competition from cable television, satellite, and the Internet, NBC and FOX decided not to broadcast the first of three scheduled presidential debates. NBC cited a legal obligation, claiming its broadcast contract with Major League Baseball required it to preempt the debate to cover the playoffs.[204] FOX offered simple economics, choosing to offer a highly

---

increase in voting following the 1960 televised debates. Voter turnout, Campbell notes, increased sharply between the 1932 and 1936 presidential elections, the same period when broadcast radio began its rapid national expansion. *Id.* By the 1960 elections, however, "90 percent of the population reported listening to radio and 80 percent read a daily newspaper." *Id.* Televised debates, therefore, merely complimented the existing political reporting, increasing the depth, but not the scope, of voter education.

200. The average national viewership for the presidential elections remained above sixty million households between 1960 and 1992. COMMISSION ON PRESIDENTIAL DEBATES, DEBATE HISTORY: 1976 DEBATES, *at* http://www.debates.org/pages/debhis76.html (last visited Mar. 20, 2002) (reporting 69.7 million viewers in 1976); COMMISSION ON PRESIDENTIAL DEBATES, DEBATE HISTORY: 1980 DEBATES, *at* http://www.debates.org/pages/debhis80.html (last visited Mar. 20, 2002) (reporting 80.6 million viewers); COMMISSION ON PRESIDENTIAL DEBATES, DEBATE HISTORY: 1984 DEBATES, *at* http://www.debates.org/pages/debhis84.html (last visited Mar. 20, 2002) (reporting 65.1 million viewers); COMMISSION ON PRESIDENTIAL DEBATES, DEBATE HISTORY: 1992 DEBATES, *at* http://www.debates.org/pages/debhis92.html (last visited Mar. 20, 2002) (reporting 62.4 million viewers).

201. COMMISSION ON PRESIDENTIAL DEBATES, DEBATE HISTORY: 1996 DEBATES, *at* http://www.debates.org/pages/debhis96.html (last visited Mar. 20, 2002) [hereinafter 1996 DEBATES].

202. Scholars have noted that televised presidential debates have "become a beacon of sanity in the electoral process," in comparison to campaign commercials. Ed Bark, *Defining Moments: Audience Will Be Watching Debates Carefully, Pros Say*, DALLAS MORNING NEWS, Oct. 2, 2000, at 1A (quoting Professor Marc Landy); *see also* Editorial, *Debates Give Voters Insight Into Election*, SAN ANTONIO EXPRESS-NEWS, Oct. 3, 2000, at 6B ("The debates are vital because they give the candidates a chance to deliver their message without being filtered by the media.").

203. Johnson v. FCC, 829 F.2d 157, 164 (D.C. Cir. 1987) (stating that televised debates are "only one of the great number of avenues for candidates to gain publicity and credibility with the citizenry"); Alexander, *supra* note 152, at 127; Spotts, *supra* note 172, at 563.

204. Howard Kurtz, *NBC Tosses Debate Choice to Affiliates; Stations Can Pick Between Politics and Playoffs*, WASH. POST, Sept. 30, 2000, at C1.

publicized action adventure premiere.[205] FCC Chairman William E. Kennard quickly issued a scathing condemnation of the networks' decisions,[206] and Commissioner Susan Ness echoed his sentiment.[207] Public criticism ranged from outrage to satire,[208] but industry executives defended the move as a simple business decision consistent with the demonstrated interests of the marketplace.[209] More tellingly, the viewers tuned out the debates and turned on the alternatives. While the baseball game failed to draw solid ratings, FOX's *Dark Angel* premiere packed in more than seventeen million households, easily beating CBS's debate coverage and nearly topping ABC's debate coverage as well.[210] Although the combined network and cable viewership ultimately demonstrated a significant national interest in the elections,[211] the week-end ratings showed that America's political appetite was largely confined to *The West Wing*.[212]

The 2000 presidential debates reveal two important dimensions to the modern American voter. First, as industry pundits have recognized, viewers seek programming alternatives. Given the option to choose professional sports, Hollywood hype, or presidential candidates, many network viewers opted out of the debates. Second, despite the poor network showing, the first debate reached more homes than either of the presidential debates held during the 1996 general elections.[213] The low network viewership masked a larger audience watching the debates on cable television channels and premium satellite stations.[214] Viewers,

---

205. Don Kaplan, *Sexy Angel Sinks Debate: Titanic Creator Launches Ratings Winner*, N.Y. POST, Oct. 5, 2000, at 94.

206. William E. Kennard, Editorial, *Fox and NBC Renege on a Debt*, N.Y. TIMES, Oct. 3, 2000, at A27.

207. Press Release, FCC Commissioner Susan Ness Decries Decisions of NBC and FOX Networks not to Air the First Presidential Debate (Sept. 29, 2000) (on file with author).

208. Stephen Hess, *Reschedule This Pesky Election*, USA TODAY, Oct. 2, 2000, at 8A (arguing that the presidential elections should be held in February, "between the Super Bowl and the NCAA [basketball] tournament").

209. *See, e.g.*, Kaplan, *supra* note 205 (quoting television analyst Marc Berman, who stated "[a]nytime you have political programming—even the presidential debates—... the other networks will benefit .... It happens every time, and it made very good sense for Fox ....").

210. *Id.*

211. Lisa de Moraes, *The Real Loser on Debate Night: NBC's Baseball Strikes Out*, WASH. POST, Oct. 5, 2000, at C7. The first debate reached approximately 46.6 million viewers, exceeding the total for the first presidential debate in the 1996 general election. *Id.*

212. *The Week's TV Ratings*, S.F. CHRON., Oct. 11, 2000, at C4 (reporting NBC's fictional series *The West Wing* ranked first in viewership for the week of the first presidential debate).

213. de Moraes, *supra* note 211 (reporting that viewership for the first debate of the 2000 general election averaged 46.6 million households); 1996 DEBATES, *supra* note 201 (reporting that viewership for the first debate of the 1996 general election averaged 46.1 million households, and 36.3 million households for the second debate).

214. *Compare* Kaplan, *supra* note 205 (using network totals to predict that total debate viewership would be less than 35 million), *with* de Moraes, *supra* note 211 (reporting actual

therefore, demonstrated a preference for both debate alternatives (content) and network alternatives (forum) for the candidates' speech.[215] These conclusions suggest that any regulation compelling the live unified broadcast of the debates must carefully consider the relevant market for debate coverage, an analysis addressed in Part III.

III. OUT OF THE CHAOS: A MARKET APPROACH TO THE CONSTITUTIONAL ANALYSIS OF BROADCAST REGULATIONS

The Supreme Court's decisions in *NBC*, *Red Lion*, and *Tornillo*, and the economic criticism of the scarcity rationale are debates about content. *Red Lion*'s deferential review and *Tornillo*'s intense scrutiny are both sufficient to manage broadcast traffic. Criticism of the public interest doctrine in broadcasting arises from the Supreme Court's First Amendment non-broadcast jurisprudence, which has long assumed that the First Amendment's core values are most prohibitive of government regulations based on content.[216] Although criticism of broadcast regulation often reaches the system of federal licensing itself, it is the content-based restrictions permitted in broadcasting, premised on the scarcity theory, that draw the greatest fire.

A. A "Quick-Look":[217] The Determinative Role of Content in First Amendment Analysis

The concern over content-based restrictions is deeply rooted in the decisions of the Supreme Court. Content-based restrictions, the Court has explained, seek to differentiate speakers solely on the basis of their perspectives, views, or beliefs.[218] Content-based restraints raise numerous conflicts with First Amendment values, distorting public debate toward a government-favored position and fostering a paternalistic intolerance for speech not sanctioned by the

---

viewership of 46.6 million).

215. Thomas W. Hazlett, *Digitizing "Must-Carry" Under* Turner Broadcasting v. FCC, 8 SUP. CT. ECON. REV. 141, 186 (2000). Professor Hazlett notes that during the 2000 election, all but one of the eighteen debates held during the presidential primaries and general election were carried live on national cable television networks. *Id.* at 186-87.

216. *See, e.g.*, Police Dep't of Chi. v. Mosley, 408 U.S. 92, 95-96 (1972); *see also* R.A.V. v. City of St. Paul, 505 U.S. 377, 382 (1992) (holding that content-based restrictions are "presumptively invalid").

217. The determinative role of content in the speech cases is analogous to the "quick-look" doctrine in antitrust law. The Supreme Court has recognized that an abbreviated economic analysis, known as a "quick look" is appropriate in cases where "an observer with even a rudimentary understanding of economics could conclude that the arrangements in question would have an anticompetitive effect . . . ." Cal. Dental Ass'n v. FTC, 526 U.S. 756, 770 (1999); *see also* Jay P. Yancey, Comment, *Is the Quick Look Too Quick?: Potential Problems with the Quick Look Analysis of Antitrust Litigation*, 44 U. KAN. L. REV. 671 (1996).

218. Helgi Walker, *Communications Media and the First Amendment: A Viewpoint-Neutral FCC is Not Too Much to Ask For*, 53 FED. COMM. L.J. 5, 6 (2000).

state.[219] The specter of government controlled speech taints the First Amendment's role in the "search for political truth"[220] by encouraging one public viewpoint at the expense of all others.[221] Therefore, assuming the type of content restrained is of sufficient value,[222] the First Amendment provides a near-absolute shield against government regulation outside the broadcast industry.[223]

In contrast, content-neutral restraints may be upheld where the government demonstrates that the regulation effectively promotes a substantial interest unrelated to viewpoint suppression.[224] Some commentators still view content-neutral restraints as potential threats to public debate capable of limiting access to sources of information and thereby skewing the discourse towards a single result.[225]

Not all commentators have accepted the Court's corollary content doctrines, which allow for reduced judicial scrutiny where the government regulates in a neutral manner without regard to the speaker's viewpoint.[226] Moreover, by

---

219. Geoffrey R. Stone, *Content-Neutral Restrictions*, 54 U. CHI. L. REV. 46, 55-57 (1987).

220. Consol. Edison Co. of N.Y., Inc. v. Pub. Serv. Comm'n of N.Y., 447 U.S. 530, 538 (1980) ("To allow a government the choice of permissible subjects for public debate would be to allow that government control over the search for political truth.").

221. Buckley v. Valeo, 424 U.S. 1, 48-49 (1976).

222. Geoffrey R. Stone, *Content Regulation and the First Amendment*, 25 WM. & MARY L. REV. 189, 194-95 (1983) (discussing Chaplinsky v. New Hampshire, 315 U.S. 568 (1942)). In *Chaplinsky*, the Court held that certain types of speech are considered to have low social value, and thus are only provided minimal constitutional protection. *Chaplinsky*, 315 U.S. at 571-72.

223. Police Dep't of Chi. v. Mosley, 408 U.S. 92, 95 (1972).

224. Schad v. Borough of Mount Ephraim, 452 U.S. 61, 70 (1981); United States v. O'Brien, 391 U.S. 367, 377 (1968). Although content-neutral restraints must limit their incidental restrictions on speech, the regulation need not be the least restrictive method of achieving the government's goal. Ward v. Rock Against Racism, 491 U.S. 781, 799 (1989).

225. Stone, *supra* note 219, at 55.

226. Critics have noted that the Court's stated reasons for strictly scrutinizing content-based restrictions are logically applicable to content-neutral restraints. Professor Martin Redish has argued that while content-based restrictions can undermine the democratic process by impeding voter education, content-neutral regulations will likely have the same effect. Martin H. Redish, *The Content Restriction in First Amendment Analysis*, 34 STAN. L. REV. 113, 128 (1981). In addition, he notes that requiring speech restraints to target all information without regard to content ultimately "reduces the sum total of information or opinion disseminated." *Id.* Similarly, Professor Erwin Chemerinsky argued that the Court has used the content-neutral exception to uphold restrictions on speech that adopt a favored viewpoint, even if neutrally applied. Erwin Chemerinsky, *Content Neutrality as a Central Problem of Freedom of Speech: Problems in the Supreme Court's Application*, 74 S. CAL. L. REV. 49 (2000). Professor Chemerinsky discussed, for example, the Court's decision in *Forbes* holding that minor party candidates for political office may be excluded from broadcast debates. *Id.* at 56-57. The Court found that the exclusion of minor party candidates was a content-neutral restriction, based on the likely success of the candidate, and not the candidate's views. Arkansas Educ. Television Comm'n v. Forbes, 523 U.S. 666, 682 (1998). Professor Chemerinsky argues that the distinction between major and minor candidates,

removing certain classes of speech such as "fighting words"[227] or obscenity[228] from First Amendment protection, the Court itself engages in an explicit content analysis.[229] Despite the criticism, however, the content doctrine has been consistently reaffirmed by the Court and is unlikely to be abandoned.

The broadcast cases depart from the content model of the First Amendment, causing the doctrinal tension between print and electronic media.[230] Content-based restrictions, the Court reasoned in *NBC*, are essential for broadcast regulation because the selection of broadcasters on anything other than a lottery system requires a content choice.[231] While a lottery system managed through property rights and capital was possible, the *Red Lion* Court feared the threat of private information monopoly.[232] If content-based decisions were essential, the *Metro Broadcasting* Court concluded, the choice should at least serve the socially beneficial purposes of "public interest, convenience, or necessity."[233] Forged in the era of national socialism,[234] the Court's content-based broadcast doctrine was thus born.

B. A Reasonable Rule for the Future of Broadcasting Analysis

Identifying content-based restraints as the problem with the broadcast cases does not, however, help select among the proposals for reconciling *NBC*, *Red Lion*, and *Tornillo*. Critics of the scarcity theory have called for a direct overruling of *Red Lion*, leading to a single broadcast standard under the holding of *Tornillo*.[235] This proposal, however, ignores the Supreme Court's concern

---

however, only existed because of the government-imposed evaluation of the public interest in each candidate's views. Chemerinsky, *supra*, at 59-60.

227. *E.g.*, R.A.V. v. City of St. Paul, 505 U.S. 377 (1992).

228. *E.g.*, Miller v. California, 413 U.S. 15 (1973).

229. *See, e.g.*, Young v. Am. Mini Theatres, 427 U.S. 50, 66 (1976) (plurality opinion stating that the First Amendment's protection "often depends on the content of the speech").

230. It should be noted that this departure applies to regulations of the broadcast industry structure, and not to regulations aimed directly at broadcast content. *See* Fox Television Stations, Inc. v. FCC, 280 F.3d 1027, 1046 (D.C. Cir. 2002).

231. NBC v. United States, 319 U.S. 190, 216-17 (1943). The *NBC* Court reasoned that

> If the criterion of "public interest" were limited to such matters, how could the Commission choose between two applicants for the same facilities, each of whom is financially and technically qualified to operate a station? Since the very inception of federal regulation by radio, comparative considerations as to the services to be rendered have governed the application of the standard of "public interest, convenience, or necessity."

Id.

232. Red Lion Broad. Co. v. FCC, 395 U.S. 367, 391, 400-01 (1969).

233. Metro Broad., Inc. v. FCC, 497 U.S. 547, 567 (1990).

234. HUBER, *supra* note 90, at 5.

235. *See, e.g.*, Telecomm. Research & Action Ctr. v. FCC, 801 F.2d 501, 509 (D.C. Cir. 1986) (predicting that "the Supreme Court will one day revisit this area of the law and . . . eliminate the

with administrative flexibility. Broadcast technologies are dynamic and changing, ill-suited to inflexible judicial standards.[236] The Court has implicitly acknowledged that Congress has passed the issue of broadcast management to the FCC and the courts with practically no guidance.[237] The Court has thus expressed that both the fact-finding powers of Congress and the daily involvement of administrative agencies in media management are important resources that should be consulted in defining the accepted doctrinal limits of broadcast speech.[238]

Each of these aspects of the current system of broadcast regulation provides limits on proposals for reform. Outright abolishment of the FCC might be consistent with normal free-market economics[239] but is highly unlikely in the near future.[240] The modern FCC is a massive bureaucracy, comprised of more than 2000 full-time employees serving in twenty-nine divisions.[241] The agency's maze of administrative, technical, and support responsibilities requires an annual budget of more than $200 million.[242] Although preserving a bureaucracy because of its size is hardly laudable, the Court is unlikely to dismantle an agency of this scope by removing the scarcity underpinning in one ruling. Congress is similarly unlikely to abolish the FCC because the agency's regulation of indecency and obscenity is too easily exploited during elections.[243] Finally, the growth of new media itself argues for at least a limited federal regulatory presence if only to order and direct the growing amounts of communication traffic.[244]

---

distinction between print and broadcast media, surely by pronouncing *Tornillo* applicable to both . . .").

236. CBS v. Democratic Nat'l Comm., 412 U.S. 94, 102 (1973) ("[S]olutions adequate a decade ago are not necessarily so now, and those acceptable today may well be outmoded 10 years hence.").

237. *Id.* at 103 ("[W]hen we face a complex problem with many hard questions and few easy answers we do well to pay careful attention to how the other branches of Government have addressed the same problem.").

238. Turner Broad. Sys., Inc. v. FCC, 512 U.S. 622, 665-66 (1994) ("Congress is far better equipped than the judiciary to 'amass and evaluate the vast amounts of data' bearing upon an issue as complex and dynamic" as cable broadcasting (quoting Walters v. Nat'l Assoc. of Radiation Survivors, 473 U.S. 305, 331 n.12 (1985))).

239. LAWRENCE LESSIG, CODE AND OTHER LAWS OF CYBERSPACE 182-85 (1999).

240. Joseph D. Kearney, *Will the FCC Go the Way of the ICC?*, 71 U. COLO. L. REV. 1153 (2000) (discussing alternatives to abolishing the FCC, including a congressional reduction of the agency's authority, or self-reduction by the FCC itself).

241. *Id.*

242. HUBER, *supra* note 90, at 5.

243. Henry Goldberg & Michael Couzens, *"Peculiar Characteristics": An Analysis of the First Amendment Implications of Broadcast Regulation*, 31 FED. COMM. L.J. 1, 42 (1978). Goldberg and Couzens stated the question bluntly, arguing that the role of the First Amendment in broadcasting is "not a question of constitutional law, and probably never has been." *Id.*

244. Self-directed reorganization by the FCC itself, of course, remains possible. *See* Roger

A possible solution, however, is presented by the Court's holdings in *Turner I* and *Reno*. In both cases, the Court revealed that the scarcity doctrine is not applicable to all media.[245] The decisions suggest that similarly plentiful communications media, such as wireless, 3G,[246] and DBS, should also be subject to traditional First Amendment scrutiny. Having cited the virtues of convergence,[247] the Court's broadcast doctrine may now be on a technological timetable that will use the arrival of broadband services to mark the close of the scarcity era.[248]

C. Moving Broadcasting Back to the Marketplace of Ideas: Using Broadcast Market Power to Determine First Amendment Scrutiny

Although convergence theory should ultimately underlie the Supreme Court's review of broadcast regulation, the scarcity doctrine remains the current standard of constitutional analysis. While the broadcast networks will continue to question their public interest duties, some issues of public concern are likely to trigger new regulatory efforts. Network broadcast coverage of the general presidential debates, for instance, is an important social interest and a potentially popular political target. The 2000 presidential race prompted immediate calls for reform of all aspects of the election process.[249] Television received particular attention, largely due to the broadcast networks' practice of projecting the winner of each state.[250] Overhauling the American voting system, however, is a

---

M. Golden, *Gauging Michael Powell*, LEGAL TIMES, May 30, 2001, *available at* http://www.law.com (last visited April 1, 2002) (quoting FCC Chairman Michael Powell's desire for reform: "[W]e are in the process of systematically reviewing and thinking through what is the optimal, organizational model for the commission").

245. *See supra* notes 101-08 and accompanying text.

246. *See* FCC, THIRD GENERATION ("3G") WIRELESS, *available at* http://www.fcc.gov/3G (last visited Apr. 15, 2002). 3G systems use radio frequencies to provide Internet, multimedia, and voice communications to wireless and mobile receivers. *Id.*

247. *See supra* note 84.

248. Nicholas Negroponte summarized the past and future of a converged broadcast media:

> In analog days, the spectrum allocation part of the FCC's job was much easier. It could point to different parts of the spectrum and say: this is television, that is radio, this is cellular telephony, etc. Each chunk of spectrum was a specific communications or broadcast medium with its own transmission characteristics and anomalies, and with a very specific purpose in mind. But in a digital world, these differences blur or, in some case, vanish: they are all bits. They may be radio bits, TV bits, or marine communication bits, but they are all bits nonetheless, subject to the same commingling and multi-use that define multimedia.

NICHOLAS NEGROPONTE, BEING DIGITAL 54 (1995).

249. Katherine Q. Seelye, *Nation Awash in Ideas for Changing Voting*, N.Y. TIMES, Jan. 28, 2001, at I12.

250. Katharine Q. Seelye, *Congress Plans Study of Voting Processes and TV Coverage*, N.Y. TIMES, Feb. 9, 2001, at A20.

complicated and politically treacherous task.[251] As swift reform appears unlikely,[252] FCC action involving the broadcast network debates presents an attractive alternative.

A regulation of presidential debate coverage on the broadcast networks would force the Supreme Court to confront the scarcity doctrine directly. Although the majority opinion in *Turner I* suggests that the Court is open to reform,[253] long-standing decisions such as *NBC* and *Red Lion* are particularly likely to command adherence from the proponents of *stare decisis*. Moving beyond scarcity before the arrival of convergence thus requires an approach that combines the administrative deference of *Red Lion* with the recognition of emerging market alternatives to broadcasting noted in *Turner I* and *Reno*.

A suitable alternative may exist in the Supreme Court's antitrust decisions. In the area of antitrust law, the Court has recognized that the once strict categorical analysis of potentially anti-competitive actions has been replaced by a more searching inquiry into the harms resulting from the restraint.[254] Similarly, under the First Amendment, avoiding the strict scrutiny applied to content-based restraints does not guarantee constitutionality, but merely subjects the regulation to something less than the "most exacting level of First Amendment scrutiny."[255] Evaluating broadcast speech likewise requires "an enquiry meet for the case, looking to the circumstances, details, and logic of a restraint."[256] As the Court reiterated in *Turner I*, the special interests permitting broadcast regulation do not

---

251. Editorial, *Election Reform Stalls*, N.Y. TIMES, Apr. 30, 2001, at A18.

252. Katharine Q. Seelye, *Voting System Changes Lag, Experts on Elections Warn*, N.Y. TIMES, Apr. 4, 2001, at A18; Katharine Q. Seelye, *Little Change Forecast for Election Process*, N.Y. TIMES, Apr. 26, 2001, at A14.

253. *See* Turner Broad. Sys., Inc. v. FCC, 512 U.S. 622, 638 & n.5 (1994) (noting that "courts and commentators have criticized the scarcity rationale since its inception"). A more direct assault on scarcity is found in Justice Blackmun's concurrence in *CBS v. Democratic National Committee*, 412 U.S. 94, 158 n.8 (1973). Justice Blackmun noted that scarcity "may soon be a constraint of the past, thus obviating the concerns expressed in *Red Lion*." *Id.* (Blackmun, J., concurring).

254. *See* Cal. Dental Ass'n v. FTC, 526 U.S. 756, 779 (1999).

255. *See Turner Broad. Sys., Inc.*, 512 U.S. at 661. This lesser or intermediate standard derives from the oft-quoted Supreme Court decision in *United States v. O'Brien*, which permits content-neutral restraints furthering an important government interest unrelated to speech suppression, narrowly tailored to limit incidental speech restraints. 391 U.S. 367, 377 (1968). *See also* Ward v. Rock Against Racism, 491 U.S. 781, 799-800 (1989) (explaining the application of narrowly tailored restraints).

256. *Cal. Dental*, 526 U.S. at 781. Justice Souter's explanation of this standard in antitrust law appears readily applicable to broadcast speech restrictions:

> The object is to see whether the experience of the market has been so clear, or necessarily will be, that a confident conclusion about the principal tendency of a restriction will follow from a quick (or at least quicker) look, in place of a more sedulous one. And of course what we see may vary over time, if . . . analyses in case after case reach identical conclusions.

Id.

"readily translate" into other communication markets.[257] A narrower focus on the specific broadcast markets restrained by a regulation would provide the flexibility to accommodate new technical innovations without deregulating the entire broadcast industry in a single step.[258]

Accordingly, the Supreme Court should adopt the market analysis that guides the evaluation of monopolization cases under the Sherman Antitrust Act[259] as the standard for reviewing speech restrictions on broadcast television networks. Speech restraints in markets regarded by First Amendment precedent as scarce, such as broadcast network television and broadcast radio, would be evaluated under the reduced First Amendment scrutiny articulated in *NBC* and *Red Lion*. In contrast, restraints in markets that are regarded by precedent as abundant, such as cable television and the Internet, would be evaluated under strict or intermediate scrutiny, depending on whether the regulation is content-based.[260] Where the regulated content is found in scarce and abundant media, the level of constitutional protection, and thus the level of scrutiny, will depend on which market carries the majority of the speech at issue. Courts would determine the "primary market" for the content by using the market power tests employed in antitrust cases.[261]

A market power[262] approach to broadcast regulation has significant advantages over the current First Amendment tests. A market approach adds the full protection of the First Amendment to speech primarily carried in media that lack the distinctive characteristics of the electromagnetic spectrum.[263] Regulations on speech found primarily in media with the distinct characteristic of spectrum scarcity[264] can still be deferentially reviewed to allow narrow federal

---

257. *See Turner Broad. Sys., Inc.*, 512 U.S. at 639 (quoting Bolger v. Youngs Drug Prods. Corp., 463 U.S. 60, 74 (1983)).

258. Flexible regulations are critical in this area, as "technological advances have a habit of moving more rapidly than government policy." R. Michael Senkowski et al., *Broadband: Flying Blind*, LEGAL TIMES, May 14, 2001, at 33 (noting the "Internet has emerged as a center of commerce, news, and entertainment in the relatively brief span since enactment of the Telecommunications Act of 1996").

259. 15 U.S.C. § 2 (1994).

260. *See supra* notes 218-25 and accompanying text.

261. *See, e.g.*, U.S. DEP'T OF JUSTICE AND THE FTC, HORIZONTAL MERGER GUIDELINES (1992) (establishing definitions for determining the amount of market power a firm possesses for a specific product) [hereinafter DOJ GUIDELINES].

262. LAWRENCE A. SULLIVAN & WARREN S. GRIMES, THE LAW OF ANTITRUST: AN INTEGRATED HANDBOOK 22 (2000). In antitrust economics, market power is defined as "the seller's ability to raise and sustain a price increase without losing so many sales that it must rescind the increase." *Id.*

263. FCC v. League of Women Voters of Cal., 468 U.S. 364, 376 (1984) (noting broadcasting regulations "involve unique considerations," justifying a departure from the First Amendment protections provided to other forums).

264. *Id.* at 377 ("The fundamental distinguishing characteristic of the new medium of broadcasting that, in our view, has required some adjustment in First Amendment analysis is that

intervention for social programs.[265] A market-based standard for determining the scrutiny of broadcast restraints begins to realign the First Amendment protections of broadcasters with all other media and gradually removes the government's role in content choice.[266] Most importantly, a market-based analysis provides broadcasters with maximum First Amendment protection over content distributed through multiple media outlets. Broadcasters are thus given a clear incentive to speed the convergence of media through broadband technologies,[267] a goal already mandated by Congress in the Telecommunications Act of 1996.[268]

Ample guidelines for this analysis already exist because the principles of

---

'broadcast frequencies are a scarce resource [that] must be portioned out among applicants'." (quoting CBS v. Democratic Nat'l Comm., 412 U.S. 94, 101 (1973))).

265. *CBS*, 412 U.S. at 157-58 (Blackmun, J., concurring). Justice Blackmun argued that the "Commission has a duty to encourage a multitude of voices but only in a limited way, viz.: by preventing monopolistic practices and by promoting technological developments that will open up new channels." *Id.*

266. The Supreme Court has stated that eliminating content-based restrictions is the central purpose of the First Amendment. Consol. Edison Co. of N.Y., Inc. v. Pub. Serv. Comm'n of N.Y., 447 U.S. 530, 538 (1980).

267. Although deployment of broadband has moved slowly, Kornbluh, *supra* note 88, United States Internet users are showing new interest in high-speed Internet capacity, as "consumers are switching from dial-up to broadband faster than new households are getting dial-up." Saul Hansell, *Can AOL Keep Its Subscribers in a New World of Broadband?*, N.Y. TIMES, July 29, 2002, at C1; *see also* Jim Hu, *More Consumers Hooked on Broadband*, CNET News (Jan. 15, 2003), *at* http://news.com.com/2100-1033-980737.html (reporting a fifty-nine percent increase in broadband use in the United States in 2002).

268. *In re* Deployment, *supra* note 89, at 20,915 (discussing Telecommunications Act of 1996, Pub. L. No. 104, § 706, 110 Stat 56 (1996)). In 2001, broadband deployment received fresh legislative attention in a host of bills in the House and Senate. The bills shared a common format, hoping to entice regional telecommunications carriers to hasten broadband implementation by exempting high-speed services from the provisions of the 1934 Act. *See* Broadband Internet Access Act of 2001, H.R. 267, 107th Cong. (2001); Internet Freedom-Broadband Deployment Act of 2001, H.R. 1542, 107th Cong. (2001); Broadband Competition and Incentives Act of 2001, H.R. 1697, 107th Cong. (2001); American Broadband Competition Act of 2001, H.R. 1698, 107th Cong. (2001); Broadband Internet Access Act of 2001, S. 88, 107th Cong. (2001); Broadband Deployment and Competition Enhancement Act of 2001, S. 1126, 107th Cong. (2001). In 2002, the House approved a measure aimed at speeding broadband development, by lifting restrictions that prohibit phone companies from offering high-speed Internet access "without first opening their local markets and permitting smaller rivals to lease their equipment." Stephan Labaton, *Broadband Bill Advances, But Its Survival is Doubtful*, N.Y. Times, Feb. 28, 2002, at C4 (discussing the Internet Freedom-Broadband Deployment Act of 2001, H.R. 1542, 107th Cong. (2001)). FCC Commissioner Kevin Martin has similarly signaled that the FCC intends to move quickly in the coming years to accelerate broadband deployment. Stephen Lawson, *FCC Commissioner Calls for Quick Decisions*, *available at* http://www.itworld.com/Man/2697/ IDG020124fccsupernet (last visited Jan. 24, 2002).

market power are well developed in antitrust law.[269] Market power is measured by determining the relevant geographic and product markets for a particular good or service.[270] The geographic market is the region in which consumers can reasonably seek alternatives to the product or service in question.[271] The product market includes all goods or services that are reasonably interchangeable with the product in question.[272] Although elasticity will normally locate substitute products or regions of competition,[273] the Supreme Court has held that in some instances a single product brand can comprise the entire relevant market.[274] Therefore, market analysis seeks to find whether a seller possesses sufficient power over a marketplace to reduce the output of supply and trigger price increases above the normal competitive level. If consumers can readily shift their consumption to competing markets without great additional expense,[275] the two markets are considered the relevant area of competition.[276]

These basic parameters can be applied to determine the First Amendment scrutiny of a broadcast regulation restraining speech in both scarce and abundant markets.[277] In step one, the geographic market for the regulation is determined by evaluating the "area" in which consumers can reasonably access alternatives to the broadcast medium restrained. For instance, a decision by the FCC denying the application of a licensee to erect a radio tower[278] is largely limited to the surrounding few miles around the proposed transmitter. As consumers are unlikely to travel to distant communities for a similar radio broadcast, the geographic market would likely be drawn narrowly. In contrast, a regulation

---

269. The foundations of market power measurement date back at least to Judge Learned Hand's opinion in *United States v. Aluminum Co. of America*, 148 F.2d 416 (2d Cir. 1945).

270. Tampa Elec. Co. v. Nashville Coal Co., 365 U.S. 320, 327-29 (1961) (defining the relevant market as the "area of effective competition").

271. United States v. Grinnell Corp., 384 U.S. 563, 574 (1966).

272. United States v. E. I. du Pont de Nemours & Co., 351 U.S. 377, 395 (1956).

273. Elasticity measures a seller's market power as the percentage of decline in demand for the seller's product in response to an increase in the price of the product. SULLIVAN & GRIMES, *supra* note 262, at 22-23. Where a seller lacks market power, an increase in price will cause buyers to stop purchasing the seller's product, denoting an elastic market. *Id.* Where the seller holds significant market power, buyer demand will not significantly decline in response to price increases, signaling an inelastic market. *Id.* For a discussion of the economic models of elasticity, see Gregory J. Werden, *Demand Elasticities in Antitrust Analysis*, 66 ANTITRUST L.J. 363 (1998).

274. Eastman Kodak Co. v. Image Technical Servs., Inc., 504 U.S. 451, 481-82 (1992).

275. *Id.*

276. DOJ GUIDELINES, *supra* note 261, at 65.

277. It is crucial to note that the following examples are based on presumptions concerning consumer behavior in situations arising in several First Amendment broadcast decisions. In antitrust cases, the definition of the relevant market is an issue of expert economic opinion and cannot normally be determined by laypersons. *See id.* (discussing the modeling of a hypothetical marketplace).

278. *See* FCC v. Pottsville Broad. Co., 309 U.S. 134, 139-40 (1940); FCC v. Sanders Bros. Radio Station, 309 U.S. 470, 471 (1940).

excluding certain candidates for political office from participating in a debate held on a state public television channel[279] impacts a broader market. As consumers here may access the broadcast debate throughout the state, the geographic market would extend to at least the state borders.

In other cases, the geographic broadcast market might be national. Regulations specifying the type of programming that may be broadcast among affiliated radio stations[280] or a generalized public service requirement such as the fairness doctrine[281] affect consumers throughout the United States.[282] Finally, a regulation similar to the Communications Decency Act, which prohibited offensive transmissions over the Internet,[283] controls a virtually unlimited geographic market.

Step two of the analysis determines the marketplace for the broadcast product regulated. Selecting the relevant product requires determining the content subject to the restriction, a more complicated problem than geography. As in antitrust analysis, a court must consider both the content that is directly regulated and any competing content that is "reasonably interchangeable."

For instance the regulation requiring cable providers to carry local broadcast television addressed in *Turner I* assumed that local television broadcasters were essential sources of information and entertainment.[284] The majority, however, found insufficient evidence that local broadcasters would be harmed without access to the cable television subscribers.[285] From a market perspective, the Court concluded that the relevant product market for broadcast information and entertainment might not be limited to local television.[286] It is important to recall that products need not be identical "or even perfect substitutes" to occupy the same product market.[287] Dissimilarities between traditional broadcast content and new media alternatives should not necessarily preclude the use of a broadcast product market that encompasses both. While some broadcast products might consist of a single outlet, others might span the spectrum of modern communications.

At step three, the market power of the relevant product in the relevant geographical area is quantified. Precise indicators of market power will vary by

---

279. *See* Ark. Educ. Television Comm'n v. Forbes, 523 U.S. 666, 671-72 (1998).

280. *See* NBC v. United States, 319 U.S. 190, 194 n.1 (1943).

281. *See supra* notes 131-44 and accompanying text.

282. A narrower geographic market might be present where the broadcast entity operated in only a specific number of cities. *See, e.g.*, FCC v. League of Women Voters of Cal., 468 U.S. 364, 370 (1984) (examining claims brought by the Pacifica Foundation, a nonprofit radio corporation broadcasting in five metropolitan markets).

283. *See* Reno v. ACLU, 521 U.S. 844, 850 (1997).

284. Turner Broad. Sys., Inc. v. FCC, 512 U.S. 622, 663-64 (1994).

285. *Id.* at 664-68.

286. *Id.* at 663 ("[C]able and other technologies have ushered in alternatives to broadcast television.").

287. ROGER D. BLAIR & DAVID L. KASERMAN, ANTITRUST ECONOMICS 108 (1985).

context, but some guidelines are possible.[288] Control of only thirty-three percent of a broadcast market should likely be insufficient to show market power, as the majority of consumers are able to access alternative content.[289] Control of ninety percent of a market, in contrast, demonstrates a lack of content alternatives and market dominance.[290] Control of a slight majority of the market estimated at sixty-four percent might suggest a decline in broadcast diversity, and depending on the context, market power.[291] In a First Amendment context, however, the ultimate question to be resolved is whether the majority of broadcast market power resides in an electromagnetic spectrum. Only regulations primarily restraining broadcasting in the scarce media are evaluated under the relaxed First Amendment standards of *NBC* and *Red Lion*. Market power residing in all other media markets demonstrates consumer alternatives outside the electromagnetic spectrum and forecloses any application of the scarcity doctrine. These economically abundant media retain full First Amendment protection and are considered under the traditional content-based distinction discussed above.[292]

This basic sketch of a First Amendment market analysis demonstrates that the law of antitrust economics provides a sound foundation for evaluating broadcast restraints. Indeed, scholars have previously demonstrated that there is nothing logically inconsistent between antitrust and free speech rationales in the area of broadcast restraints on program content,[293] and market power concerns underscore the speech issues confronted in *Turner I*.[294] Although application of this model will undoubtedly vary with facts and context, the basic reasoning is simple enough to be codified by the FCC and manageable enough for routine judicial application. In Section IV, this market model is tested against a new regulation, which compels broadcast television networks to provide live coverage of the general presidential election debates.

---

288. Although *United States v. Aluminum Co. of America*, 148 F.2d 416 (2d Cir. 1945), is regarded as the leading judicial opinion on market power, most courts now rely on the more detailed economic balancing of the DOJ Merger Guidelines. The *Alcoa* formula, using market shares of ninety percent, sixty-four percent, and thirty-three percent, is thus included only as one possible standard.

289. *Id.* at 424.

290. *Id.* at 425.

291. *Id.* at 424.

292. *See supra* notes 218-25 and accompanying text.

293. Owen M. Fiss, *The Censorship of Television*, 93 NW. U. L. REV. 1215, 1228 (1999). Professor Fiss states that in general, "a highly competitive industry is a step toward freedom insofar as it proliferates sources of information." *Id.*

294. *Id.* at 1228-29. Professor Fiss notes that on one level, the must-carry provisions at issue in *Turner I* and a subsequent case, *Turner Broadcasting System, Inc v. FCC*, 520 U.S. 180 (1997) (*Turner II*), attempted to preserve competition in the television industry by ensuring that broadcast television could access the cable television market, and thus the cable television audience. Fiss, *supra* note 293, at 1228-29.

IV. REGULATING THE DEBATE MARKET: EVALUATING THE CONSTITUTIONALITY OF COMPELLED NETWORK DEBATE COVERAGE USING A MARKET-BASED FIRST AMENDMENT THEORY

A regulation compelling broadcast television networks to cover the general presidential election debates is a useful example for applying a market-based First Amendment analysis. The presidential debates combine the two concepts cited by the Supreme Court in *Turner I* as most likely to warrant restriction of broadcast speech. First, the presidential debates have historically aired on network television, the medium the Court described as a "principal source of information . . . for a great part of the Nation's population."[295] Second, the debates serve to educate the voting public, a goal that *Turner I* implied was "a governmental purpose of the highest order."[296] A regulation designed around these considerations would force the Court to consider a speech restriction on content that served a concededly important public interest, partially aired in the forum that is still firmly controlled by caselaw decided on the basis of spectrum scarcity.

A. Some Suggested Goals for a Broadcast Debate Regulation

It is difficult and unnecessary to speculate on the precise language of a possible debate regulation. Certain provisions, however, are likely to be essential. These provisions are not an exhaustive list or a minimum set of requirements.[297] Instead, these guidelines reflect the current format of the presidential debates and some recurring problems in their coverage:

1. Unified Coverage: Traditionally, all four of the major television networks have aired the general presidential debates. This unified coverage is necessary to reduce viewer attrition,[298] which appears to occur when even minor programming alternatives are offered.[299] Technological alternatives, such as offering only "split-screen" coverage of the debates, would be similarly prohibited.
2. Live Coverage: Political commentary has become a media staple. Within minutes of the final question, analysts descend on the airwaves with evaluations, criticisms, and of course, the announcement of the victor. While commentary can serve a legitimate journalistic function, viewing the recap

---

295. Turner Broad. Sys., Inc. v. FCC, 512 U.S. 622, 663 (1994).

296. *See id.* (discussing public affairs and educational programming in general).

297. Further, these recommendations do not address more substantive matters, such as question selection, choice of moderator, or format.

298. *See* CNN, *Fourteen Million Opt for "Dark Angel" Over Debate*, Oct. 5, 2000, *at* http://asia.cnn.com/2000/ALLPOLITICS/stories/10/04/debate.tv.ap (last visited Apr. 7, 2002) (quoting Dan Rather's characterization of the first 2000 presidential debate as "pedantic, dull, unimaginative, lackluster, humdrum—you pick the words").

299. *See supra* note 185 and accompanying text; *see also* NEWTON N. MINOW ET AL., PRESIDENTIAL TELEVISION 8 (1973) (explaining the impact of a politician "appearing simultaneously on most major television channels, so that alternative viewing choices are sharply limited . . .").

before viewing the candidates risks alienating the audience. Airing the debates as they occur thus ensures that substance precedes spin.[300]

3. Full Length Coverage: The value of the debates is the opportunity to examine a candidate's responses to a wide range of issues. Airing only a portion of the presidential debate necessarily involves selecting which of the topics covered is sufficiently trivial for preemption.
4. Running Time: More debate is not necessarily better debate. The presidential debates have generally run between one and one and one half hours.[301] The one-hour running length is a reasonable standard that avoids losing viewers as the debate progresses.
5. Free Television: Although declining in recent years, broadcast television continues to account for more than fifty percent of prime time viewing.[302] In order to reach this majority of viewers, network debate coverage would be limited to the free television spectrum rather than a network-owned cable channel.
6. Commercial Sponsorship: Although network sponsorship of the debates is little more than a historic accident,[303] declining outside sponsors avoids the intrusive interruption of commercials. While sponsorship could be arranged without commercial pause, the unseemly sight of a corporate icon hovering over the nation's next leader is inconsistent with the importance of the election. If necessary at all,[304] advertisements should be limited to short

---

300. The concept of live television itself, however, may soon become antiquated with the rise of personal video recorders, or "PVRs." PVRs operate like traditional videocassette recorders, using high-capacity hard disk drives for storage in place of magnetic tape. David Pogue, *State of the Art; Recorders to let you Tame TV*, N.Y. TIMES, Apr. 5, 2001, at G1. PVRs provide a digital "buffer" between the broadcast signal and the viewer by storing up to thirty minutes of programming as it airs. *See* http://www.tivo.com (last visited Jan. 1, 2002). PVRs are thus able to customize even live television, allowing viewers to skip or re-watch segments as desired.

More importantly, PVRs, in conjunction with digital cable or satellite television services, allow viewers to choose programs weeks before they are broadcast. Pogue, *supra*. While home recorders are not new, PVRs add the element of "time-shifting," as owners pre-select their viewing by content, and not broadcast time. As one commentator writes, "[y]ou'll never know or care when a particular program was on, or on what channel; you will just know that when the little light on the front of the PVR is on, something you requested is ready to play." *Id.* As the popularity of PVRs grows, the preemptive power of unified broadcasting will vanish, as viewers will be able to watch a disk full of their favorite programming rather than the live offerings scheduled for a given timeslot. PVRs are now installed in an estimated one million homes, with future sales expected to reach fifteen million within five years. *In re* Annual Assessment, *supra* note 80, para. 94.

301. *See generally* COMMISSION ON PRESIDENTIAL DEBATES: DEBATE HISTORY, *available at* www.debates.org (last visited Apr. 30, 2002) (listing running times for all televised presidential debates).

302. *In re* Annual Assessment, *supra* note 80, para. 80.

303. *See supra* note 166 and accompanying text.

304. Prohibiting advertising might not be unreasonable given that broadcasters still "profit immensely from election campaigns." THOMAS E. PATTERSON, THE VANISHING VOTER 175 (2002).

segments no more than a few minutes at the beginning and close of each debate.[305]

7. Preemption of Sports: Perhaps most importantly, the model regulation must address the numerous network contracts with professional sports franchises. Sporting events are typically broadcast under long-term and highly profitable exclusive contracts. A model regulation must therefore supersede the networks' obligations under these contracts by exempting performance.[306]

Each suggestion seeks to minimize the networks' financial hardships during the broadcasts, while preserving the educational benefits of minimal programming alternatives.

B. The Test Applied: Determining the Relevant Broadcast Market for the General Presidential Debates

Predicting market competition without careful economic analysis risks public policy choices that stifle market growth and yield inefficient regulation.[307] Abstract broadcast market scrutiny is equally difficult given the constant changes in technology. The First Amendment scrutiny applied to a broadcast debate regulation will thus depend largely on the circumstances surrounding its enactment. Where are the viewers? What are people watching? How do televisions, computers, and even radios receive information? Each answer depends entirely on how we divide the broadcast spectrum in the future.[308]

In 2000, for instance, candidates for public office spent more than one billion dollars on television advertising. *Id.*

305. JOHN ELLIS, NINE SUNDAYS: A PROPOSAL FOR BETTER PRESIDENTIAL CAMPAIGN COVERAGE 26 (Joan Shorenstein Barone Center on the Press, Politics and Public Policy, John F. Kennedy School of Government, Harvard University 1991). This innovative proposal suggested a system of ninety-minute prime-time debates each Sunday for nine weeks, rotating among the major networks and independent news stations. *Id.* at 4. The proposal recognized the need to induce networks to sponsor the debates by permitting limited advertising sales: "if handled correctly, commercials should not diminish the value of the broadcast. Viewers and voters are sophisticated enough to understand the need for commercial sponsorship." *Id.* at 26. *But see* PATTERSON, *supra* note 304, at 174. Professor Patterson argues that the "Nine Sundays" plan is unduly burdensome, and proposes a less ambitious alternative requiring the networks to devote a single prime-time hour to each candidate for an interview hosted by the network's news anchor. *Id.* at 173-74.

306. The proposal would also require restrictions preventing minor networks from offering substitute coverage. One possibility is to require sports franchises to agree that any preempted event would be aired on a substitute channel of the network's choosing, allowing for subcontracting to a rival network, or more likely, an in-house cable station.

307. *See generally* ROBERT H. BORK, THE ANTITRUST PARADOX: A POLICY AT WAR WITH ITSELF (1978).

308. *See* Hazlett, *supra* note 77, at 927 (noting that any definition of broadcast technology can be altered by "further subdivision of time, power, or bandwidth coordinates"); *see also* LESSIG,

For the present, therefore, we are limited to the model of a broadcast television debate designed in 1960. This model, discussed above in the guidelines for a suggested regulation,[309] allows a short inquiry into the First Amendment restrictions that may govern actual promulgation.

The market-based approach begins by defining the market for the general presidential debates geographically and as a product. Defining the relevant geographic market is a straightforward task, as the proposed regulation addresses only the presidential election, rather than contests for state-specific offices. Accordingly, a national geographic market is appropriate. Next, the possible product market is defined, beginning narrowly[310] and assuming that broadcast network television is the relevant medium. A market limited to only broadcast television could be too narrow, avoiding television's overlap with other media. Excerpted transcripts of the debates, for instance, are commonly published in national newspapers[311] and Internet databases.[312] Viewers may also watch the debates on the Internet,[313] cable television,[314] or premium DBS services.[315] Non-broadcast resources, however, lag behind as widely used alternatives for debate audiences.[316] While access to Internet and cable has exploded, studies evidence a digital divide that limits the spread of information technologies in low-income

---

supra note 239, at 184 (describing broadcast technologies modeled on computer networks that avoid overlapping signals and allow endless amounts of content); Lee Gomes, *Boomtown: Visionaries See a Day When Radio Spectrum Isn't Scarce Commodity*, WALL ST. J., Sep. 30, 2002, at B1 (discussing "open spectrum" technologies).

309. *See supra* notes 297-306 and accompanying text.

310. DOJ GUIDELINES, *supra* note 261 (explaining that the relevant market model begins with "the smallest group of products" that might satisfy consumer demand).

311. *See, e.g.*, *The 2000 Campaign: Exchanges Between the Candidates in the Third Presidential Debate*, N.Y. TIMES, Oct. 18, 2000, at A26; *The 2000 Campaign: Second Presidential Debate Between Gov. Bush and Vice President Gore*, N.Y. TIMES, Oct. 12, 2000, at A22; *The 2000 Campaign: Transcript of Debate Between Vice President Gore and Governor Bush*, N.Y. TIMES, Oct. 4, 2000, at A30.

312. *See, e.g.*, C-Span.org, Presidential Debates 2000, *at* http://www.c-span.org/campaign2000/presdebates.asp (last visited Apr. 27, 2002) (archiving all of the presidential and vice-presidential debates from the 2000 election, along with commentary, discussion boards, and general election statistics).

313. *See* 1996 DEBATES, *supra* note 201 (containing downloadable video of the 1996 presidential debates). IBM and Sony Electronics plan to convert 115,000 hours of video produced by CNN since 1980 into a computerized database. Susan Stellin, *CNN Video Archives to Become Digital Database*, N.Y. TIMES, Apr. 23, 2001, at C8. The system will allow "the sale of news video material to the public on a pay-per-view basis on the Internet or through high-speed interactive cable systems." *Id.*

314. *See supra* note 214 and accompanying text.

315. Press Release, EchoStar Communications Corp., EchoStar's DISH Satellite Television Offers 504 Hours of Free Air Time to U.S. Presidential Candidates (Oct. 2, 2000) (on file with author).

316. *See supra* note 214 and accompanying text.

and rural regions.[317] Broadcast television, at present,[318] remains the most prevalent medium for debate access throughout the nation.[319] Without evidence of more widespread consumer use of new media, the initial product market is limited to broadcast television.

The market-based approach must also consider the general presidential debates as a product. The market again begins narrowly, including only the live debates before considering reasonably interchangeable debate alternatives. Considering the presidential debates a separate product market might ignore consumer habits. Media studies suggest that the debate audience is largely comprised of viewers who closely follow all developments in the election.[320] The debates may also be interchangeable with campaign advertisements, live rallies, stump speech coverage, or the candidates' web sites.[321] Yet the presidential debates—however marginalized by appeals to showmanship—remain unique in their ability to convey both the style and substance of the candidates.[322] Viewers watching the debates merely to reinforce their initial candidate choice are still held captive in front of competing viewpoints easily skimmed over in print.[323] And image does matter. The visual presentation of the candidates without protection from staff members or the safety of a teleprompter, offers insight into a candidate's ability to think clearly and respond decisively.[324] Even with

---

317. *See* William E. Kennard, *The Digital Divide*, *at* www.fcc.gov/commissioners/kennard/col051298.html (last visited Apr. 15, 2002); ESA, NATION ONLINE, *supra* note 83, at 11-29 (outlining demographic factors in computer and Internet usage in the United States). Economic barriers to new technology are referred to as "switching costs" and are recognized as a limitation on consumer alternatives sufficient to narrow the relevant market to a single product or area. Eastman Kodak Co. v. Image Technical Servs., Inc., 504 U.S. 451, 481-82 (1992).

318. One legislative initiative in 2001 sought to encourage broadband development in rural areas. Rural Broadband Deployment Act of 2001, S. 1127, 107th Cong. (2001). The Rural Broadband Deployment Act would exempt carriers providing advanced telecommunications services in areas with a "population less than 50,000 located outside of a metropolitan statistical area," from the Communications Act of 1934. *Id.* The bill defined advanced telecommunications services as the "capability to transmit information at no less than 384 kilobits per second in at least one direction." *Id.*

319. PATTERSON, *supra* note 304, at 175. Professor Patterson notes that the flood of viewing options offered by new media might increase the relative importance of the networks because "[a]s the number of channels grows, viewers stop surfing and limit their search to selected ones." *Id.*

320. *See* Campbell, *supra* note 161.

321. Both presidential candidates maintained their own campaign web sites during the election. Archived copies are available at http://www.georgewbush.com and http://www.algore2000.com.

322. Alexander, *supra* note 152, at 125-26.

323. LEMERT ET AL., *supra* note 198.

324. JAMIESON & BIRDSELL, *supra* note 154, at 15 (discussing the "often maligned but nonetheless important characteristic of . . . image"); *see also* MINOW ET AL., *supra* note 299, at 4 ("Not speeches on the stump, not speeches from the rear platform of trains, not courthouse square handshaking, not newspapers, not magazines, not books, and not even radio can confront so many people with the president's face and with his words at the moment he utters them.").

advances in technology, therefore, the live presidential debate should remain its own product market.[325]

Using this market-based analysis, a model presidential debate regulation should be viewed as limited to the live candidate debates broadcast nationally on network television. The First Amendment scrutiny applied to the regulation then depends on the amount of content regulated in this relevant market. Using the total viewership for the first general presidential debate of the 2000 campaign as a benchmark, approximately thirty-five million households of the 46.6 million households watching relied on broadcast television coverage.[326] Broadcast television thus carried more than seventy-five percent of the presidential debate audience, a market share strongly suggesting market power within the scarce broadcast medium.[327] With viewership concentrated within the traditional television spectrum, the relaxed First Amendment standards of *NBC* and *Red Lion* should be applied to review any regulation of the presidential debates.

These relaxed standards triggered by the market-based First Amendment analysis suggest that the FCC could compel the major broadcast television networks to cover the general presidential debates. In turn, the public interest goals of voter education and informed election discourse are likely sufficient regulatory concerns to pass the deferential First Amendment review required under precedent.

A broadcast debate regulation would not, of course, survive constitutional challenges indefinitely. As noted, cable television accounted for roughly twenty-five percent of the households tuning in for the first presidential debate of 2000. If convergence keeps pace, the major broadcast networks will offer but one of the many options for viewing campaign debates. In the meantime, the relaxed First Amendment scrutiny applied to a current debate regulation ensures continued network television coverage and preserves a national voter audience for emerging media alternatives.

CONCLUSION

As government control of broadcast speech approaches its seventy-fifth year, two pressing problems have emerged. Broadcast networks face mounting competition from communications media, decreasing their willingness to perform public interest duties assigned by the FCC. What network television invented in 1960, the presidential debate, it may dismantle by means of defection before

---

325. *See* PATTERSON, *supra* note 304, at 164. Professor Patterson argues that the drama of the live debates uniquely satisfies the political interests of the average voter, because "[v]oters are not like students in a classroom," but "more like the crowd at a ball game." *Id.* at 164-65. Accordingly, "[t]he more exciting the game, the more attention spectators pay. And the more attention they pay, the more they understand what's happening on the field." *Id.* at 165.

326. *Supra* note 211.

327. *See supra* notes 288-91 and accompanying text (assuming that a market share greater than sixty-four percent likely demonstrates market power).

2004.[328] The presidential debates are neither perfect nor essential to American democracy. But they are an important part of our political tradition, adding a symbolic, and sometimes substantive, focus to the selection of our highest office.[329] Simultaneously, the same technologies that have triggered competition in the communications industry are quickly eroding the already doubtful scientific basis of the FCC's most powerful regulatory schemes. Today, if not in 1927, broadcast media are not scarce.

Politics makes legislative solutions difficult to craft: no member of Congress is eager to voice support for an end to regulation of broadcast obscenity and media indecency. Administrative solutions are promised,[330] but the sheer size and power of the FCC make change difficult.[331] While the courts remain hesitant to intrude, reform is possible within the normal confines of judicial review.

A market-based approach to the First Amendment rights of broadcasters is a sensible, familiar alternative to the current two-tiered system of constitutional review. A market-based approach to the First Amendment adds a sophisticated set of guidelines suitable for agencies, broadcasters, and courts. Market-based First Amendment rights preserve the traditional deference to agency regulation in broadcast television and radio. At the same time, emerging technologies are accorded the robust speech protection of the common law First Amendment. New media are given the freedom to flourish, while old media are given a reason to catch up.

A regulation compelling the broadcast coverage of presidential debates is a fitting forum to welcome the new First Amendment rights of broadcasters. While remnants of scarcity concerns still control, federal oversight of the debates is appropriate. As the broadcast marketplace of ideas begins to rely on the economic market, the legal and scientific gap that separates Twentieth Century

---

328. Post-election news coverage of presidential politics supports this trend. On November 8, 2001, President George W. Bush delivered his second televised address to the nation regarding the United States' war on terrorism. Despite the obvious importance of the event, only ABC carried the speech live. *Bush Loses in Network Battle of "Survivor", available at* http://www.cnn.com/2001/showbiz/TV/11/08/networks.snub.bush/index.html (last visited Nov. 8, 2001). NBC and CBS each decided to air their popular prime time properties "Friends," and "Survivor," while FOX left the programming decision to the local affiliates. *Id.*

329. The importance of televised debates is gradually emerging in other countries as well. *See* Steven Erlanger, *German Candidates Unscathed After First Televised Duel*, N.Y. TIMES, Aug. 26, 2002, at A3.

330. Roger M. Golden, *Gauging Michael Powell: What Can Business Expect from a New FCC Chairman Promising Change?*, LEGAL TIMES, May 18, 2001, at 32. Chairman Powell stated that "[w]ith increasingly converged services, it is difficult to rationally label and, thus, assign regulatory treatment to an innovative provider, product or service." *Id.*

331. Yochi J. Dreazen, *FCC's Powell Quickly Marks Agency as His Own*, WALL ST. J., May 1, 2001, at A28. Chairman Powell has criticized the FCC's public interest doctrine as "about as empty a vessel as you can accord a regulatory agency and ask it to make meaningful judgments." *Id.* In contrast, FCC Commissioner Gloria Tristani saw no ambiguity in the agency's duty, citing "70 years of good, clear case law about the public interest standard." *Id.*

jurisprudence from Twenty-first Century technology can be crossed. Compelled debate broadcasts, like scarcity, must ultimately give way to the reality of a converged media. That future will validate the First Amendment's core commitment to public debate and usher in a new era of digital democracy.

Indiana Law Review

Volume 36 **2003** **Number 1**

NOTES

CREATING AN UNCOMFORTABLE FIT IN APPLYING THE ADA TO PROFESSIONAL SPORTS

JEFFREY MICHAEL CROMER*

> Fact is, every day your body feels a little different and golf is such a finite game that a little off can translate into a lot. One or two degrees here and there can mean from four to seven yards. That's not a whole lot but it's magnified due to the precision the game demands.[1]

INTRODUCTION

In May 2001, the Supreme Court had the opportunity to determine how several provisions of the Americans with Disabilities Act (ADA),[2] specifically "public accommodation," "fundamental alteration," and "private entity" should fit into athletic competition when it decided *PGA Tour, Inc. v. Martin*.[3] Instead, the Supreme Court imposed the ADA on professional sports organizations without appreciating the basics of the game of golf—or the law. The Supreme Court's expansive interpretations of "reasonable accommodation" and "fundamental alteration" under the ADA, as well as the Court's assumption that nine reclusive jurists could decide what constitutes a fundamental alteration to a professional sport, have potential implications on how the laws governing disabled Americans relate to competitive sports at all levels of society, from mere recreation and exercise to professional athletics. These potential implications cannot be what Congress intended when passing the bill in 1990.

This Note examines Casey Martin's case against the PGA Tour as a foundation for analyzing the application of the ADA to professional sports, arguing that courts must distinguish competitive sports from forms of recreation and exercise when determining fundamental alterations. *Martin* is distinguishable from typical reasonable accommodation and fundamental alteration cases. In professional golf, as in all professional sports, the essence of the activity is competition.[4] When a governing authority is forced to alter an

---

* J.D. Candidate, 2003, Indiana University School of Law—Indianapolis; B.A., 1998, Wabash College, Crawfordsville, Indiana.

1. TIGER WOODS, TIGER WOODS: HOW I PLAY GOLF 9 (2001).
2. 42 U.S.C. § 12101 (1994).
3. 532 U.S. 661 (2001).
4. Brief for the ATP Tour, Inc. and the Ladies Professional Golf Ass'n, Amici Curiae at 3,

existing rule of a sport in order to accommodate the disabled, that change, no matter how small, will often result in a fundamental alteration of the event. This rule change, mandated by a court unfamiliar with the rules of the sport and the role they play in competition, creates a scenario where other participants are at a competitive disadvantage due to the waiver of a rule for a single participant.

When the Court fails to consider competition as an integral part of the sport, as opposed to mere recreation or exercise, the ADA ceases to be fair to fellow competitors. The purpose of the ADA, after all, is to provide *equal access, not equal opportunities to win*. Admittedly, there are circumstances when the ADA may be applied in competitive settings: the ADA will never allow a complete denial of access to a sport. Access is either denied to an individual or it is not, and if it is denied, the ADA applies. The question the courts must face becomes how to properly *weigh* the other side of the equation: the impact on the affected activity—in this instance, professional golf. Courts must attempt to balance the nature of the impairment on the disabled player as against the nature of the impairment on the quality of play and the nature of the sport—whether there is a fundamental alteration. Courts must be sensitive to the potential implications a reasonable accommodation may create in competitive settings, where if applied to only one individual, it may affect the nature of play to the detriment of all competitors.

This Note is not meant to discredit the purpose of the ADA and the many successes it has had in providing disabled Americans opportunities they would otherwise not have had. Nor is the purpose of this Note to argue that the ADA should never be applied to professional sports. Instead, this Note uses Casey Martin's three-year battle against the PGA Tour as a foundation to promote two separate arguments. First, under the provisions of the ADA, the PGA Tour does not qualify as a place of public accommodation. The goal of the ADA is commendable, but extending it to private organizations and clubs like the PGA Tour is not.[5] The Supreme Court must distinguish between providing mere accessibility on the public golf course and changing the rules and nature of the sport's highest level of competition.[6] The public golf course operator can and should expect to make reasonable accommodations for those disabled and desiring access to the sport.[7] However, an otherwise private organization should not change its rules simply because it operates a course for a very limited period

---

PGA Tour, Inc. v. Martin, 532 U.S. 661 (2001) (No. 00-24), noting that

> [a]thletic competition is based on nothing more than an agreed-upon set of uniform rules, all of which, in interplay with one another, define the game. At bottom, the singularity of a sports competition and the autonomy of its rules compel the conclusion that any court-imposed modification or abrogation of a rule of how the game is played changes its fundamental nature. Such fundamental alteration is not required by the ADA.

5. Nikos Leverenz, *Americans with Disabilities Act Must Not Be Used to Change the Rules of Professional Sports*, at http://www.pacificlegal.org.

6. *Id.*

7. *Id.*

of time.[8]

Second, the Court failed to distinguish adequately the applicability of the ADA to competitive settings from its applicability to recreational exercise when it determined what constitutes a fundamental alteration under the ADA. Specifically, the Supreme Court failed to recognize that the nature of a professional golf event is based solely on competition. By applying the ADA's reasonable accommodation requirement to the PGA Tour, courts will inadvertently erode the fundamental fairness of competition and alter the sport in a way only those competitors directly involved can recognize. An elite group of golfers has the "total package" to play at a level worthy of earning millions of dollars and travel throughout the world to compete.[9] Professional play takes place at a vastly more elevated level than that of recreational golfers. Only professional golfers know what is needed and what changes affect their level of competition. Only their sport's governing body has the expertise and ability to dictate the rules and truly recognize what is essential to a sport, and determine what "fundamentally alters" an event.

For these reasons, this Note discusses why the ADA does not comfortably fit in professional sports and explores possible implications of *Martin*'s expansive interpretation in other professional and competitive sport settings. Part I of this Note first analyzes the history and purpose of the ADA, specifically Title III, which is most applicable to *Martin* and professional sports settings.[10] This analysis gives the proper background to explain why the PGA Tour should not have been considered a public accommodation under the ADA. Part II explains the story of Casey Martin and sets out the legal history of his case against the PGA Tour. Part III of this Note examines how courts have historically applied the ADA in competitive sports settings. Part IV discusses the recent decision of the Supreme Court, examining the strengths and oversights of the majority and dissent.

After this background of the ADA and the Supreme Court's decision in *Martin*, Part V of this Note examines and critiques the Supreme Court's expansive interpretation in three respects: 1) the scope of "reasonable accommodation" as it applies to the PGA Tour and other professional sports organizations, 2) what constitutes a "fundamental alteration" in a professional

---

8. *Id.*

9. Brief for the ATP Tour, Inc. and the Ladies Professional Golf Ass'n, Amici Curiae at 10, *Martin* (No. 00-24), stating,

> the fact that Martin was substantially disadvantaged by application of this uniform rule is precisely the point: athletes who, for whatever reason, cannot perform as well as others when measured against the uniform rules of the game are not supposed to win! In fact, if their particular physical traits (e.g., height, weight, strength, endurance) or psychological characteristics (e.g., ability to concentrate, risk aversion, "mental toughness") create sufficient barriers, they may not be able to compete effectively at all. The "game" is designed to test for those characteristics and to make them outcome determinative, not to compensate for individual differences.

10. *See* 42 U.S.C. § 12182(a) (1995).

sports setting, and 3) what body is in the best position to determine what constitutes a fundamental alteration. This section includes a discussion of why such decisions should be entrusted to the sport's own governing body for reasons of institutional competence. Finally, Part VI discusses the potential implications of the Supreme Court's decision in *Martin* on professional sports organizations and other levels of athletic competition. This part also discusses the Supreme Court's application of the ADA in a recent case and its attempt to potentially narrow the reach of the ADA in all areas of law, including athletic competitions.

The ADA does not fit as comfortably into competitive settings as into traditional employment and recreational applications. As Tiger Woods observed, a small change may not seem like much to the spectator, but every change is magnified for the player due to the precision of the game,[11] often resulting in a fundamental alteration that only those familiar with the sport can sense.[12] Athletics played at the sport's highest level are clearly different than sports played for recreation, exercise, or as a hobby.[13] Professional sports are games of precision and are decided by the slightest of margins. The Court should recognize this distinction and apply the ADA accordingly.

I. BACKGROUND OF THE ADA

A. General Purpose

Congress enacted the Americans with Disabilities Act[14] in 1990 to "assure equality of opportunity."[15] The purpose of the ADA was to "provide a clear and comprehensive national mandate for the elimination of discrimination against individuals with disabilities."[16] The ADA prohibits discrimination based on disability. Unlike the ADA, other employment discrimination laws forbid employers to take into account a particular disability of an employee,[17] while the

---

11. Brief for Petitioner at 37 n.25, PGA Tour, Inc. v. Casey Martin, 532 U.S. 661 (2001) (No. 00-24) (noting "the effect need not be great to have a significant impact on the competition. In 1997, for example, the difference in average score between the number one and number 100 PGA TOUR scoring leaders was 2.32 strokes per round over the course of the entire year.").

12. WOODS, *supra* note 1.

13. Brief for the ATP Tour, Inc. and the Ladies Professional Golf Ass'n, Amici Curiae at 3, *Martin,* (No. 00-24) (noting "[a]thletic competition is supposed to favor the more skilled and physically able. It is a test of who is the 'best' at mastering the game as defined by its rules, and it is this characteristic that makes it compelling to both competitors and spectators.").

14. *See* 42 U.S.C. §12101.

15. Alex B. Long, *A Good Walk Spoiled: Casey Martin and the ADA's Reasonable Accommodation Requirement in Competitive Settings*, 77 OR. L. REV. 1337, 1342 (1998) (quoting 136 Cong. Rec. S9694 (daily ed. July 13, 1990)).

16. Martin v. PGA Tour, Inc., 994 F. Supp. 1242, 1247 (D. Or. 1998) (quoting 42 U.S.C. § 12101(b)(1) (1995)).

17. *See* Title VII of the Civil Rights Act of 1964, 42 U.S.C. §§ 2000e (1994); *see also* Age Discrimination in Employment Act, 29 U.S.C. §§ 621-34 (1994).

ADA expressly requires employers to make reasonable accommodations for disabled employees.[18]

Title I of the ADA pertains to discrimination in the workplace.[19] Since its passage in 1990, Title I of the ADA has impacted labor law, preventing covered entities from discriminating against a disabled individual in the areas of job hiring, advancement, and other facets of employment.[20]

"Title II is targeted at public entities, making it unlawful for them to discriminate against or exclude qualified, disabled individuals from participating in or receiving the benefits of their services, programs, or activities."[21] The definition section provides that Title II is meant for those "who, with or without reasonable modifications to rules, policies, or practices, the removal of architectural, communication, or transportation barriers, or the provision of auxiliary aids and services, meets the essential eligibility requirements for the receipt of services or the participation in programs or activities provided by a public entity."[22] Title III applies to private entities providing places of public accommodation,[23] and Title IV relates to telecommunications.[24] Title IV applies to the Federal Communications Commission, and addresses the retaliation, coercion, state immunity and discrimination in that area.[25]

1. Title III.—Title III is the most relevant section of the ADA for the Martin case and other situations involving disabled Americans and competitive settings. Title III constitutes "perhaps the most ambitious section of the ADA, granting rights to disabled customers who would not otherwise have been permitted to participate in the central activities of mainstream society."[26] Title III prohibits discrimination against disabled persons in the "full and equal enjoyment of the goods, services, facilities, privileges, advantages, or accommodations of any place of public accommodation by or any person who owns, leases (or leases to), or operates a place of public accommodation."[27] Private clubs are exempt from Title III provisions, a vital and often overlooked argument presented by the PGA Tour in *Martin*.[28] Discrimination is defined in these circumstances as the use of any eligibility requirement that as a result screens out disabled persons from their equal enjoyment of the public accommodation.[29] Title III also covers the failure

---

18. *See* Long, *supra* note 15, at 1343 (citing 42 U.S.C. §§ 12112(b)(5)(A), 12182(b)(2)(A)(ii)).

19. 42 U.S.C. § 12181(7).

20. Christopher M. Parent, *A Misapplication of the Americans with Disabilities Act*, 26 J. LEGIS. 123, 129 (2000).

21. *Id.* at 131.

22. *Id.*

23. 42 U.S.C. § 12182(a) (2002).

24. *Id.* § 12201-12213.

25. *Id.*

26. Parent, *supra* note 20, at 131.

27. 42 U.S.C. § 12182(a) (2002).

28. *Id.* § 12187.

29. *Id.* § 12182(b)(2)(a)(I).

to make reasonable accommodations to the disabled person in situations that are necessary to provide services and goods to the disabled individual.[30] The only defense by employers or clubs under Title III, which was an argument by PGA Tour in the case of *Martin*, is if they can demonstrate that such a modification would "fundamentally alter" the nature of such goods or services.[31]

A "private entity" under Title III is "any entity other than a public entity."[32] A "public entity" is "any State or local government; any department, agency, special purpose district, or other instrumentality of a State or States or local government."[33] If that were the extent of what constitutes a public entity, there would not have been a question as to whether the PGA Tour must accommodate Martin. Instead, immediately following the public entity section, Title III lists various private entities that the ADA will consider public for purposes of providing accommodations to disabled Americans. This list contains, among other places, hotels, restaurants, movie theatres and stadiums.[34] This list concludes by expressly naming golf courses as private entities that the ADA considers a public entity.[35]

Titles I and III of the ADA also state that private entities must make reasonable modifications "unless the entity can demonstrate that taking such steps would fundamentally alter the nature of the good, service, facility, privilege, advantage, or accommodation being offered or would result in an undue burden."[36] A public entity's claim of fundamental alteration or undue burden is the only defense a public entity can claim from the requirements of the ADA. Is walking a golf course a fundamental alteration to the game? Arguably, yes. More importantly, should a non-profit, private entity, which merely uses a public golf course for four days a year and prohibits the public from access to the playing areas, be considered a public entity? The Supreme Court found so, to the surprise of many spectators of the sport and the Court alike.

II. GOLF AND LEGAL HISTORY OF CASEY MARTIN

On May 29, 2001, Casey Martin prevailed over the PGA Tour when the U.S. Supreme Court ruled that Martin was entitled to use a golf cart while playing on

---

30. *Id.* § 12182(b)(2)(a)(ii).

31. *Id.*

32. 42 U.S.C. § 12181(6).

33. *Id.* § 12131(1)(A),(B).

34. Scotta A. Weinberg, *Analysis of* Martin v. Professional Golvers' Ass'n, Inc.—*Applying the ADA to the PGA is a Hole in One for Disabled Golfer*, 38 BRANDEIS L.J. 757 (2000) (citing 42 U.S.C. § 12181(7) (2002)).

35. *Id.* (citing 42 U.S.C. § 12181(7)(L)).

36. 42 U.S.C. § 12182(2)(A)(iii); Parent, *supra* note 20, at 132 (noting that "Titles I and III are closely linked as they demand many of the same requirements on the covered entity, primarily that 'reasonable accommodations' be made and that any eligibility criteria intended to screen out an individual with a disability are eliminated.").

the PGA Tour.[37] *PGA Tour, Inc. v. Martin* was the first case in which a party asked the Supreme Court to apply the ADA to a competitor in a professional sports organization.

Casey Martin suffers from Klippel-Trenaunay-Weber Syndrome, a rare vascular congenital defect that causes a lack of blood circulation leading to bone deterioration and muscle atrophy wherever the disease is found.[38] Martin suffers from this disease in his right leg, which causes severe pain and discomfort while walking.[39] Martin's defect has progressed to the point at which it is unsafe for him to walk for any significant period of time without the possibility of fracturing his leg.[40]

Casey Martin can hit golf balls well enough to be part of an elite group of players throughout the world able to earn a living playing golf at its highest level. There is no dispute Martin is a very talented golfer and has the ability to compete with other professional athletes on the PGA Tour. However, Martin's degenerative leg condition substantially limits the amount of walking he is capable of doing, especially the amount of walking necessary to compete on the PGA Tour. An average PGA Tour event normally involves walking twenty to twenty-five miles over the four day period.[41] When Martin played collegiate golf at Stanford University, he was noticeably in pain from walking according to his competitors.[42] Even his opposition begged their coaches to lift the mandatory-walking rule so that Martin could ride a cart.[43] The coaches allowed Martin to ride a golf cart while his competitors walked the course, and the National Collegiate Athletic Association agreed with the waiver.[44] Upon winning the 1994 NCAA Championship with Stanford, Martin pursued a career as a professional golfer at the highest level of competition, the PGA Tour.[45]

With a chance to play on the PGA Tour, Martin decided to challenge the PGA Tour's rule requiring all competitors to walk the course during all tournaments. Martin requested to use a golf cart during a qualifying tournament and the PGA Tour denied his request.[46] Martin sought and was granted a

---

37. PGA Tour, Inc. v. Martin, 532 U.S. 661 (2001).

38. Martin v. PGA Tour, Inc., 994 F. Supp. 1242, 1243 (D. Or. 1998) (summary of Casey Martin's disorder and its side effects).

39. *Id.*

40. *Id.*

41. Brief for Petitioner at 4, *Martin* (No. 00-24) (noting that "during four days of competition, a golfer typically must walk twenty to twenty-five miles, often in intense heat, in inclement weather, or over hilly terrain. The cumulative fatigue resulting from this requirement may, and frequently does, affect golfers' concentration, shot-making ability, and overall performance.").

42. Marcia Chambers, *The* Martin *Decision*, GOLF DIGEST, Aug. 2001, at 60.

43. *Id.*

44. *Martin*, 532 U.S. at 668.

45. *Id.*

46. *Id.* at 662.

preliminary injunction allowing him to use a golf cart.[47] After qualifying for the Nike Tour, a tour co-sponsored by the PGA and governed by PGA rules, the United States District Court for Oregon extended the preliminary injunction for the first two tournaments on the Nike Tour.[48] By qualifying for this PGA sponsored tour, which complies with mandatory-walking rule of the PGA Tour, Martin was able to bring an ADA-based suit against the PGA Tour.[49] Martin contended that by failing to provide him with a cart while playing, the PGA Tour failed to make its tournaments accessible to individuals with disabilities, in violation of the ADA.[50]

The PGA Tour is a non-profit entity that was formed in 1968.[51] The PGA Tour is made up of three annual tours, the PGA Tour, the Nike Tour (currently the "Buy.com Tour") and the Senior PGA Tour, which use sponsors and co-sponsors to fund their various tournaments throughout the year. Approximately 200 golfers participate on the PGA Tour, roughly 170 on the Buy.com Tour and about 100 on the Senior PGA Tour.[52] The PGA Tour operates tournaments that are mostly four-day events and played on courses leased and operated by the PGA Tour.[53] Spectators are able to purchase tickets to the tournaments, but their access is strictly limited. Only PGA Tour players, caddies, and various officials are allowed "between the ropes."

There are several ways for competitors to gain entry into the three PGA-sponsored tours.[54] If a player wins three Buy.com Tour events in the same year, or is among the top fifteen money winners on the Buy.com Tour, he earns a "PGA Tour Card."[55] This simply means that he has earned the right to play in all the tournaments provided by the PGA Tour. A golfer may obtain a spot in an individual tournament through "competing in 'open' qualifying rounds, which are conducted the week before each tournament."[56]

However, the most common method of earning a PGA Tour Card is by competing in a three stage-qualifying tournament known as "Q-School."[57] Any member of the public whose "golf handicap" is below a set standard may enter Q-School by paying a $3000 entry fee and submitting two letters of reference.[58] This fee covers the player's cost to play and use of golf carts, which are permitted during the first two stages of Q-School, but have been prohibited

---

47. PGA Tour, Inc. v. Martin, 984 F. Supp. 1320, 1327 (D. Or. 1998).
48. Parent, *supra* note 20, at 133.
49. *Id.* at 132.
50. *Id.* at 133.
51. *Martin*, 532 U.S. at 665.
52. *Id.*
53. *Id.*
54. *Id.*
55. *Id.*
56. *Id.*
57. *Id.* at 665.
58. *Id.*

during the final stage since 1997.[59] Over a thousand participants enter the first stage, and only 168 enter the final stage. From the final stage, only a quarter qualify for PGA Tour membership, while those remaining enter the Buy.com Tour.[60]

Everyone would like to see someone with Martin's talent and persistence achieve success at the highest level of his profession.[61] Putting aside the human dimension of the case, though, and basing the argument solely on the legal issues, the PGA Tour was correct in rejecting Martin's request to use a golf cart during the PGA Tour events. Though Martin could drive, chip, and putt as well as other players on the PGA Tour, he lacked the ability to walk the course, a prerequisite for all players competing at golf's highest level.[62] Whether this ability is as essential to the game as driving, putting and chipping is not the issue. Rather, in a game where the slightest alteration to the rules and a change in the way one person is allowed to play is made, the results are magnified due to the quantifiable numbers used to determine who wins. Following this logic, permitting Casey Martin to ride a golf cart during tournaments fundamentally alters the competitive nature of the game.

III. PRIOR ADA APPLICATIONS IN COMPETITIVE SETTINGS

Martin was the first opportunity for the Supreme Court to interpret some of the basic provisions of the ADA as they apply to professional sports organizations. However, several recent decisions show that courts must be careful when attempting to make a reasonable accommodation for a disabled athlete without fundamentally altering an athletic competition.[63]

In *Slaby v. Berkshire* the U.S. District Court of Maryland addressed a similar situation to that of Casey Martin.[64] Members of a golf club brought suit when the club erected rope barriers around the golf course and enacted rules limiting carts on the course. Mr. Slaby was required to use a golf cart to play golf because he had three heart operations, was diabetic, suffered from hypertension, and was classified as permanently disabled by Social Security.[65] The court held that the rope barriers were only a minor inconvenience for the golfers and did not prevent the other members from playing golf, which is all that is required under the ADA in this application.[66]

Sandison v. Michigan High School Athletic Ass'n[67] demonstrates that an

---

59. *Id.*

60. *Id.*

61. Parent, *supra* note 20, at 136 (noting Martin took part in Nike's recent golf advertising campaign with inspirational messages such as "I can" and "Anything is possible.").

62. *Id.* at 145.

63. Long, *supra* note 15, at 1359.

64. 928 F. Supp. 613 (D. Md. 1996).

65. *Id.* at 614.

66. *Id.* at 615.

67. 64 F.3d 1026, 1035 (6th Cir. 1995).

exception to an established rule of a sport organization would not only create a fundamental alteration to the nature of the event, but would constitute an undue burden on the governing body.[68] In *Sandison*, plaintiffs were two disabled high school students who were two years behind in school, leaving them nineteen years old in their senior year.[69] Both were prevented from competing in cross-country events because the Michigan High School Athletic Association (MHSAA) prohibited any student over nineteen from competing.[70] The Sixth Circuit held that the effect of allowing these students to compete would create both a fundamental alteration to the event as well as an undue administrative burden because no one, including the MHSAA, can make the determination of whether something is unfair to other competitors without creating an undue burden.[71]

Finally, perhaps the most pertinent language used in recent case law regarding the applicability of the ADA to professional sports was in *Stoutenborough v. National Football League*.[72] In *Stoutenborough*, a citizen with a hearing impairment claimed that individuals with hearing impairments were effectively prevented from listening to blacked-out football games over the radio and were denied access to the games.[73] Though the Sixth Circuit held that there was no statutory basis for the claim, it provided important language when referring to Stoutenborough's Title III claim.[74] The court found that the NFL was not covered by the ADA.[75] The Sixth Circuit found that although football games are played in a place of public accommodation and can be viewed on television in other places of public accommodation, these circumstances do not support a Title III claim.[76] Because *Stoutenborough* did not involve the ADA being applied directly to an athlete, its application is limited with regard to *Martin*; however, its language is still persuasive and applicable to *Martin*. Despite tournaments being played on a place of public accommodation and being viewed on televisions in similar places, under *Stoutenborough* the PGA Tour should not be considered a place of public accommodation.

The most important lesson from theses cases and *Martin* is the importance

---

68. Long, *supra* note 15, at 1363.

69. *Id.* at 1362.

70. *Id.*

71. *Id.*

> [I]n order to conclude that the plaintiffs' age would not provide them with an unfair advantage, coaches and physicians would have to consider the skill level of each member of an opposing team, the overall skill level of each opposing team, and the skill level of each student whom the older student displaced from the team. Therefore, permitting older students to compete would fundamentally alter the nature of the program and constitute an undue administrative burden.

72. 59 F.3d 580 (6th Cir. 1995).

73. *Id.* at 582.

74. *Id.*

75. *Id.*

76. *Id.*

of courts' awareness that the ADA's reasonable accommodation requirement within the employment (Title I) and recreational settings must be distinguished from competitive and professional sport settings in order to effectively determine whether the accommodation is creating a fundamental alteration to the event. Courts must recognize that although a change in a recreational setting may not alter the nature of the event, the same change may distort the rules and strategies of a competitive setting.

There may be occasions where the ADA will be correctly applied to athletic competitions. The ADA was intended to apply and will be applied to remove artificial, man-made barriers, such as the lack of wheelchair ramps or arbitrary physical requirements. Martin's condition that made him unable to play, however, was his disease, not neutrally applied rules. The application of the ADA in this sense is analogous to equal protection law. In *Washington v. Davis*, the Supreme Court stated "we have not held that a law, neutral on its face and serving ends otherwise within the power of government to pursue, is invalid under the Equal Protection Clause simply because it may affect a greater proportion of one race than of another."[77] Discriminatory impact alone, without proof of invidious intent, raises no constitutional issue.[78] The law must and does recognize that not all participants are similarly situated. In fact, that is the very premise of competition.

IV. SUPREME COURT'S DECISION

The PGA Tour did not argue that Casey Martin was not a person with a disability covered by the ADA.[79] Rather, its initial argument was that the PGA Tour does not constitute a "place[] of public accommodation" as defined under Title III of the ADA.[80] Secondly, the PGA Tour argued that walking the golf course is essential to the game of golf at the PGA Tour level and to allow Martin to ride a golf cart would fundamentally alter the game at that level. [81]

This is the first time the Supreme Court has been asked to apply the ADA to a professional sports setting. There is little precedent as to not only how the Court should apply the ADA to *Martin*, but perhaps more importantly, as to what role, if any, the Court should play in determining whether riding a golf cart would be a fundamental alteration.

Although the Supreme Court had not dealt with athletic competitors and the ADA prior to *Martin*, the Supreme Court had indicated a desire to interpret the ADA in a very broad manner in prior decisions.[82] In *Bragdon v. Abbott*, a dentist

---

77. 426 U.S. 229, 242 (1976).

78. *Id.*

79. Tim A. Baker, *The Law and the Links: How Casey Martin Prevailed in His Legal Battles with the PGA Tour*, RES GESTAE, Sept., 2001, at 17.

80. *Id.*

81. *Id.*

82. *Id.*

refused to treat a patient who was infected with HIV.[83] The Court concluded that HIV was an impairment from the moment of infection that substantially limited respondent's ability to reproduce, which was a major life activity.[84] The Court held that a person diagnosed with HIV is a disabled person under the ADA even though the patient's infection had not yet progressed to the symptomatic phase.[85]

Further, the Court unanimously held in *Pennsylvania Department of Corrections v. Yeskey* that state prisons and prisoners are included within the coverage of the ADA.[86] In *Yeskey*, respondent prisoner was eligible for the Pennsylvania Motivational Boot Camp under a Pennsylvania statute.[87] However, because petitioner Department of Corrections refused his admission to the program due to his hypertension, respondent brought action challenging petitioner's denial of his access to the boot camp as a violation of the ADA.[88] Petitioner claimed that the ADA did not govern it or the program.[89] The Court held that state prisons were clearly subject to the ADA, and that the boot camp was an ADA-protected voluntary program by virtue of its definition in the Pennsylvania statute.[90]

Despite the trend of a broad reading of the ADA, the Supreme Court has never applied the ADA to an organization that is, arguably, a private entity. Additionally, the Court has never delved into an area where any modification of the rules would fundamentally alter the landscape of a sporting event allowing one person to change the playing field and putting fellow competitors at an inherent disadvantage that directly affects their livelihood.

A. Brief Summary of the Majority Opinion

The Supreme Court decided two distinct issues in holding that the PGA Tour was required to allow Casey Martin to ride a golf cart: "1) whether the ADA protects access to professional golf tournaments by a qualified entrant with a disability; and 2) whether a disabled contestant may be denied the use of a golf cart because it would 'fundamentally alter the nature' of the tournaments. . . ."[91]

The PGA Tour first argued that because "Title III [of the ADA] is concerned with discrimination against 'clients and customers' seeking to obtain 'goods and services' at places of public accommodation," Martin could not bring a Title III claim because he was not such a client or customer of the PGA Tour.[92] Title III states "the term 'individual or class of individuals' refers to the clients or

---

83. 524 U.S. 624 (1998).
84. *Id.*
85. *Id.*
86. 524 U.S. 206 (1998).
87. *Id.* at 208.
88. *Id.*
89. *Id.*
90. *Id.* at 210.
91. Baker, *supra* note 79 (quoting PGA Tour, Inc., v. Martin, 532 U.S. 661, 664-65 (2001)).
92. *Id.* (citing *Martin*, 532 U.S. at 678).

customers of the covered public accommodation that enters into the contractual, licensing or other arrangement."[93] The majority held, however, that [the PGA Tour] offers the public the privilege of both watching the golf competition and competing in it. "Although the latter is more difficult and more expensive to obtain than the former, it is nonetheless a privilege that the PGA Tour makes available to members of the general public."[94]

The second issue decided by the majority was whether the use of a golf cart fundamentally alters the nature of professional golf tournaments. The majority concluded that the essence of golf has always been "shot-making."[95] The majority rejected testimony from the district court that stressed the history and tradition of walking the course in the highest level of golf competition.[96] The majority ruled that the fatigue factor involved was not a large enough factor to create a fundamental alteration of the sport.[97] Finding that "shot-making" was the essence of golf, the Court found that allowing Martin to use a cart would not fundamentally alter the game of golf.

B. Brief Summary of the Dissent

In his dissent, Justice Scalia, joined by Justice Thomas, made clear that he did not necessarily believe that Casey Martin should not be able to ride a golf cart.[98] According to Justice Scalia, the legal principles on which the majority decided the case were flawed. To begin, the majority's opinion, in Scalia's view, was not based on the legal principles, but instead were based on the Court's morals and compassion for the disabled golfer.[99]

According to Justice Scalia, "The [Americans with Disabilities Act] seeks to assure that a disabled person's disability will not deny him *equal access* to (among other things) competitive sporting events—not that his disability will not deny him *an equal chance to win* competitive sporting events."[100] Everyone should have the right to not only be able to play golf on public golf courses, but everyone should have the right to try to compete at the sport's highest level. In Justice Scalia's view, the law does not permit someone to have equal opportunity

---

93. 42 U.S.C. 12182(b)(1)(A)(iv).

94. *Martin*, 532 U.S. at 680.

95. *Id.* at 683.

96. *Id.* at 685.

97. *Id.* at 683.

98. *Id.* at 704 ("My belief that today's judgment is clearly in error should not be mistaken for a belief that the PGA TOUR clearly *ought not* allow respondent to use a golf cart. That is a close question, on which even those who compete in the PGA TOUR are apparently divided, but it is a *different question* from the one before the Court.") (Scalia, J., dissenting).

99. *Id.* at 691 ("[T]oday's opinion exercises a benevolent compassion that the law does not place within our power to impose.") (Scalia, J., dissenting); *id.* at 704 ("I have no doubt Congress did not authorize misty-eyed judicial supervision of such a revolution.") (Scalia, J., dissenting).

100. *Id.* at 703.

to win at that level and, in essence, level the playing field.[101] Competition, at its core, is created by people of different abilities. To take away that inherent competition and level the playing field, Justice Scalia concluded, is to destroy the essence of golf and all of competitive athletics.[102]

V. ANALYSIS OF SUPREME COURT'S OPINION

Casey Martin's story of overcoming adversity is inspirational for all people with or without a disability, regardless of whether one is an athlete. Unfortunately, the Supreme Court's broad interpretation of the ADA in Martin has expanded the intent of the ADA, which was to ensure that disabled Americans receive equal access to everyday activities. While pursuing golf simply as recreation or at the highest level as a professional is a right that ought to be guaranteed to all citizens, playing on the PGA Tour is not. Rather, it is a privilege that ought to be granted only to those who meet all of the necessary qualifications, and the private governing body ought to have the authority to determine what those necessary qualifications are.[103] As stated in the brief of the PGA Tour:

> Individual competitors may or may not be able to compensate for those physical disadvantages, but, if they cannot, the resulting "inequality" is not discrimination in any meaningful sense (including the sense contemplated by Title III), but simply a reflection of the varying challenges faced, to a greater or lesser degree, by all athletes in elite athletic competitions. Those competitions reward superior physical performance, without adjusting the standards from competitor to competitor to allow for more equal results.[104]

The Court did try to narrow the scope of its ruling, stating that its holding pertained specifically to Martin using a golf cart on the PGA Tour, and not the general use of carts on the PGA Tour.[105] Regardless of the Court's intent, their application of the ADA shows that "few human activities are currently beyond

---

101. *Id.* ("[T]he very nature of competitive sport is the measurement, by uniform rules, of unevenly distributed excellence.") (Scalia, J., dissenting).

102. *Id.* ("[B]y giving one or another player exemption from a rule that emphasizes his particular weakness . . . is to destroy the game.") (Scalia, J., dissenting).

103. Brief for the ATP Tour, Inc. and the Ladies Professional Golf Ass'n, Amici Curiae at 3, *Martin,* (No. 00-24), noting,

> A competition judicially managed to eliminate this or that "unfairness" may appear more "fair" in the view of a court because less skilled or less able-bodied individuals may be able to compete, but it is not the same athletic competition envisioned by the creators and fans of the game. The fundamental fairness of the game—i.e., that all the rules apply equally to all competitors—has changed, and we no longer have a competition that tests who is the "best" at that particular game.

104. Brief for Petitioner at 33-34, *Martin* (No. 00-24).

105. Leverenz, *supra* note 5.

scope of either legislative regulation or judicial incursion."[106]

Three issues arise out of the *Martin* case. The first issue, and arguably the most overlooked aspect of the decision, was the holding by the Supreme Court that the PGA Tour was not a private club for purposes of the ADA and therefore not immune from its coverage.[107] The second issue, the Court's conclusion requiring the PGA to accommodate Martin by allowing him to ride in a golf cart because it would not alter the fundamental nature of the competition, drew wide praise and criticism.[108] Golfing traditionalists and historians criticized the Supreme Court's rulings while the general public found the requirement to be the proper application of the ADA. Finally, the issue left open by the majority is what role, if any, the governing body should have in determining fundamental alterations in competitive settings, including the PGA Tour and other professional organizations. Did the Supreme Court go beyond its reach and decide an issue better left to the sport's governing body?

A. Critique of Court's Reasonable Accommodation Interpretation

The most common debate stemming from the *Martin* decision was the Court's determination that allowing the use of a cart was not a fundamental alteration of the sport. Many golf and sport enthusiasts at all levels of competition were upset at the determination that an organization designed to establish the rules of a game, an organization that has had that sole authority since its inception, could be required by the judicial system to alter those rules. On the other hand, many were inspired by Martin's persistence and agreed that Martin was a perfect example of how the ADA can serve its original purpose.

The issue of whether the PGA Tour was a place of public accommodation or a private entity is a question that the majority superficially addressed. The issue of whether riding a golf cart is a fundamental alteration will always be debated. However, the issue of whether the PGA Tour should even be held to the ADA standards is, arguably, the tougher issue for the majority to defend.

The Supreme Court established what constituted a "private club" in *Moose Lodge No. 107 v. Irvis*.[109] There, the Supreme Court created a list of factors to determine whether a club is private: 1) whether the club has well-defined requirements for membership; 2) whether the club conducts all of its activities in a privately owned building or grounds; 3) whether it is not publicly funded; and 4) whether only members and guests are permitted in the club and may become a member upon invitation.[110] Under these general guidelines and another

---

106. *Id.*

107. Long, *supra* note 15, at 1338.

108. *Id.*

109. Todd A. Hentges, *Driving in the Fairway Incurs No Penalty:* Martin v. PGA Tour, *Inc. and the Discriminatory Boundaries in the Americans with Disabilities Act*, 18 LAW & INEQ. 131, 158 (2000) (citing Moose Lodge No. 107 v. Irvis, 407 U.S. 163 (1965)).

110. *Moose Lodge No. 107*, 407 U.S. at 171.

private membership test set forth by the Ninth Circuit,[111] the PGA Tour could quite easily be considered a private club. The PGA Tour provides no services to the public and membership to the PGA Tour is based solely on the golf score a player shoots.[112] In the majority opinion, Justice Stevens focused on the Q-School factor, which allows any member of the public to enter Q-School if they pay $3000 and have two letters of reference.[113] By allowing the public access to Q-School, the PGA Tour does provide access to the public that classifies the PGA Tour as a place of public accommodation.

This reasoning by the majority does not put the issue of Q-School in the proper perspective. The majority later found that the essence of the game of golf is shot-making, and that there is restricted access to the playing areas during competition.[114] The use of Q-School to gain entry to the PGA Tour is analogous to professional baseball teams holding open tryouts to the public. These tryouts occur throughout the year and in various parts of the country. As long as an applicant falls within the age range that the specific team sets forth, anyone may try out to play professional baseball. Is this the customary way to gain access to professional baseball? Absolutely not. Just as in golf, there are several other ways to make a team and be considered a professional. Under this law, professional baseball should be held to the same standards, though it is not. Therefore, the public's participation in the Q-School hardly establishes the PGA Tour as a place of public accommodation. The access to the playing areas only by the competitors should play a more important factor in determining what constitutes a place of public accommodation.

To be clear, the PGA Tour did not claim that it was a private club altogether exempt from Title III's coverage.[115] The PGA Tour admitted that its tournaments take place at places of public accommodation.[116] This argument is worthy of elaboration, though it was not addressed by either party. Instead, the issue presented by the PGA Tour was that "the competing golfers are not members of the class protected by Title III. . . ."[117] In supporting its position that it is not a public accommodation under the ADA, the PGA Tour argued that Title III is concerned with discrimination against "clients and customers" seeking to obtain "goods and services" at places of public accommodation,[118] and Title I applies

---

111. Hentges, *supra* note 109, at 158 (citing Richard v. Friar's Club, 124 F.3d 212, 217 (9th Cir. 1997) (defining the test of a private club as whether the organization: 1) is a club in the ordinary sense of the word, 2) is private, and 3) requires meaningful conditions of limited membership)).

112. *Id.* (noting that membership in the PGA Tour is highly selective and based on criteria that is clearly quantifiable and does not allow for any type of discrimination).

113. PGA Tour, Inc. v. Martin, 532 U.S. 661, 680 (2001).

114. *Id.* at 683.

115. *Id.* at 677.

116. *Id.*

117. *Id.* at 678.

118. 42 U.S.C. 12182(b)(1)(A)(iv).

to the people who work at these public accommodations.[119] "Title III is intended to confer enforceable rights on clients and customers of places of public accommodation, not on persons working to provide those clients and customers with the relevant goods and services."[120] The majority held, however, that the PGA Tour's argument fails in this respect because it offers the public the privilege of both watching the golf competition and competing in it.[121]

This reasoning is flawed for a number of reasons. First, the ADA should not be applicable "inside the ropes" because the paying public is not allowed in any playing area. There is exclusivity between the clients and customers and the players.[122] The customers, as the public, are forced to stay "outside the ropes," clearly distinguishing clients and customers from professional golfers such as Casey Martin.[123] The PGA Tour distinguished areas of play from areas outside the boundaries of play "by noting that in a typical ballpark, the stands must be accessible to the disabled because this is where the public is allowed. On the contrary . . . dugouts are not subject to the ADA because the public is not allowed there."[124]

"Golf courses are specifically mentioned in Title III because disabled individuals should be provided with the opportunity to engage in the recreation of their choice."[125] The ADA was enacted to provide the Casey Martins of the world the opportunity to play golf at courses throughout the country.[126] The

---

119. *Id.*

120. Brief for Petitioner at 11, *Martin* (No. 00-24) (noting that respondent is not in the category of customer, "[l]ike a concert hall performer, or actor in a theatre production, respondent is helping to supply the entertainment at Tour events, not seeking to enjoy it.").

121. *Martin*, 532 U.S. at 680 ("Although the latter is more difficult and more expensive to obtain than the former, it is nonetheless a privilege that petitioner makes available to members of the general public.").

122. Brief for the ATP Tour, Inc. and the Ladies Professional Golf Ass'n, Amici Curiae at 4, *Martin* (No. 00-24), noting,

> Simply because the sections of the course set aside for the gallery are open to the public does not render the competitive area where the public is not permitted (such as the playing area of an LPGA golf tournament, or the field of a professional baseball league, or the tennis courts at the National Tennis Center during the U.S. Open, or a bob-sled track during the Olympics) subject to the public accommodations provisions of Title III.

123. *Martin*, 532 U.S. at 693 (As Justice Scalia pointed out in his dissent, Title III covers only clients and customers of places of public accommodation. First, "[t]he persons 'recreat[ing]' at a 'zoo' are presumably covered [by the ADA], but the animal handlers bringing in the latest panda are not."). *Id.* Justice Scalia continued, "To be sure, professional ballplayers participate in the games, and use the ball fields, but no one in his right mind would think that they are customers of the American League or of Yankee Stadium. They are themselves the entertainment that the customers pay to watch. And professional golfers are no different." *Id.* at 695.

124. Weinberg, *supra* note 34 (quoting Martin v. PGA Tour, Inc., 984 F. Supp. 1320, 1326-7 (D. Or. 1998)).

125. Parent, *supra* note 20, at 137.

126. *Martin*, 532 U.S. at 699. "If a shoe store wishes to sell shoes only in pairs it may; and

ADA does not provide the opportunity to compete at the highest level of a sport, governed by a private entity, so long as he can fulfill the majority of the requirements needed to abide by the rules. The ADA only states that places of public accommodation include, but are not limited to, "a gymnasium, health spa, bowling alley, golf course, or other places of exercise or recreation."[127] Similarly in *Slaby*, the court found that the golf course was not denying Slaby the right to exercise or recreate. He was allowed on the golf course as much as he wanted; the club was not infringing on his right to exercise or recreate at the club.[128]

Casey Martin was not using the golf course to exercise or recreate; rather, he was using the golf course to compete at the sport's highest level and earn a living.[129] Just because "golf courses" are specifically listed in the ADA, the Court should not automatically hold that a golf course is a place of public accommodation.[130] Instead, it should consider the purpose of use and the nature of the competition. The PGA Tour did not deny Martin the right to exercise or recreate on the golf course. It did deny him the right to compete at the highest level, which does not violate the ADA.

The PGA Tour is not itself a golf course and so is not among the items listed as public accommodations under the ADA. The PGA Tour is a membership organization that sponsors events open only to those who are qualified to participate. "Other than ensuring that particular courses used by the PGA Tour meet its strict standards and regulations during the event, the Tour cares little about how those courses operate during the rest of the year."[131] The ADA also requires an individualized inquiry to determine whether a specific accommodation for a specific person's disability would be reasonable under the specific circumstances, without fundamentally altering the nature of the sport.[132] The ADA's requirement that each individual's case be considered on its own facts and not governed by blanket rules was a factor in the Court's holding that the PGA Tour must comply with the ADA. After finding that the ADA covered

---

if a golf tour (or a golf course) wishes to provide only walk-around golf, it may. The PGA TOUR cannot deny respondent access to that game because of his disability, but it need not provide him a game different (whether in its essentials or in its details) from that offered to everyone else."

127. 42 U.S.C. § 12181 (7)(L).

128. 928 F. Supp. 613, 615 (D. Md. 1996).

129. *Martin*, 532 U.S. at 695. Justice Scalia pointed out in his dissent, Casey Martin "did not seek to 'exercise' or 'recreate' at the PGA TOUR events; he sought to make money (which is why he is called a *professional* golfer)."

130. Brief for the ATP Tour, Inc. and the Ladies Professional Golf Ass'n, Amici Curiae at 13, *Martin*, (No. 00-24), noting,

> notwithstanding that they may be specifically identified on the ADA's list of public accommodations, places that are purely private in nature are not regulated places of public accommodation. For example, although a "library" is included on the list, a private library in an individual home is not a place of public accommodation because it is not a "place of public display."

131. Parent, *supra* note 20, at 137.

132. Baker, *supra* note 79 (citing 42 U.S.C. § 12182(b)(2)(A)(ii)).

the PGA Tour and that walking the course was not fundamental to the game, the Court examined Martin's personal circumstances.[133]

The Court criticized the PGA Tour for failing to look at Martin's individual circumstances.[134] However, because the PGA Tour originally believed it was not covered by the ADA, there was no need to evaluate the specific circumstances of Casey Martin. There was no reason compelling the PGA Tour to do so under these circumstances, since it considered itself a private entity under the ADA standards.[135] To criticize the PGA Tour for overlooking Casey Martin is to overlook the valid defense raised by the Tour, exempting it from the requirements of the ADA.

B. Critique of Court's Fundamental Alteration Interpretation

As the first case to apply the ADA to professional sports, "*Martin* had the opportunity to address the shortcomings of the statute in such a setting and to help define the contours of what 'reasonable accommodation' means in a situation that produces clear winners and losers based on quantifiable performance."[136]

The most apparent flaw in the majority's decision in *Martin* was the Court's failure to fully address the inherent difference between recreational settings, and the case presented here—the highest level of competition—a professional sports organization. This distinction turns on the fact that a fundamental alteration is much more apparent and can be determined by a court more easily when presented with a general place of recreation, or a place intended for exercise, as the ADA intended.[137] However, in the case at bar and other cases involving high levels of competition, a court must be intimately familiar with the nuances of the sport and must be sensitive to any slight change in the rules that will alter the landscape of a level playing field for professional players, among whom there is so little disparity in performance.[138] Simply put, the Court is ill-equipped to decide what constitutes a "fundamental alteration."

---

133. *Martin*, 532 U.S. at 690.

134. *Id.*

135. Brief for the ATP Tour, Inc. and the Ladies Professional Golf Ass'n, Amici Curiae at 9, 10, *Martin*, (No. 00-24), noting the Ninth Circuit's failure

> to appreciate fully the essential characteristics of athletic competition is nowhere more obvious than in its conclusion that an "individualized inquiry" is necessary to determine, on a competitor-by-competitor basis, whether a relaxation or modification of a rule of the game for a disabled competitor would fundamentally alter the game, or give a particular disabled competitor an advantage.

136. Long, *supra* note 15, at 1339.

137. 42 U.S.C. § 12181(7)(L).

138. *Martin*, 532 U.S. at 699. "[R]ules are the rules. They are (as in all games) entirely arbitrary, and there is no basis on which anyone—not even the Supreme Court of the United States--can pronounce one or another of them to be 'nonessential' if the rule maker (here the PGA TOUR) deems it to be essential." (Scalia, J., dissenting).

The slightest unfairness could have "major quantifiable results for other competitors that are lacking in noncompetitive settings."[139] The ADA requires the Court, not the PGA Tour, to determine the fatigue of Casey Martin and compare that to the "normal" fatigue of walking the course. For reasons of institutional incompetence, however, the Court is not the proper referee for this call; instead, it should defer to the PGA to determine what is a fundamental alteration to this private and specialized enterprise.[140]

The essential question in determining what constitutes a fundamental alteration in golf depends on the definition of professional golf. This deceptively simple question can be debated among golfers just as passionately as it was before the Supreme Court. Professional golf consists of detail-specific rules that one must learn before he can simply go out and tee off on the PGA Tour. From rules regulating the number of clubs a player may have in a golf bag to rules stating who may carry the bag, professional golf is more than just "shot-making." A player must be familiar with the rules, strategic in his club selection, and mentally prepared just to be able to play on the PGA Tour. Regardless of one's interpretation and definition of professional golf, the sport differentiates winners and losers based on a specific number—the number of shots an individual takes throughout the tournament—which is a culmination of stamina, weight, weather, swing, equipment, and of course, skill. Recreational golf, on the other hand, is simply playing golf for enjoyment and trying to improve your game. Whether a player is riding a cart or walking, the bottom line is pleasure, not worrying about conforming with specific rules and competing for money. Recreational golf not only differs in degree from professional golf, but also in kind, where the many nuances at the professional level ultimately determine the final results.

C. Critique of the Judiciary's Role in Determining What Constitutes a Fundamental Alteration

Unless the ruling body who determines what is a fundamental alteration has the ability to differentiate a professional competitive setting, where a very slight change will result in a fundamental alteration, from a pure recreational setting, where it would take a more substantial change to constitute a fundamental alteration, simply applying the ADA to professional settings will be a misapplication of this far-reaching statute. Accordingly, the PGA Tour should be its own governing body.

The PGA considers professional golf a test of not only ball striking but also competition where the element of physical stress and fatigue are added with the addition of the walking rule.[141] The PGA believed "the overall purpose of Title

---

139. Long, *supra* note 15, at 1378.

140. *Id.* (noting that the difficulty in dealing with the reasonable accommodation requirement of the ADA in competitive settings is that "the ADA essentially requires a court to measure an unquantifiable factor 'the level of Casey Martin's fatigue vs. the fatigue of able-bodied golfers' in a program based on quantification (professional tournament golf).").

141. Brief for Petitioner at 4, *Martin* (No. 00-24), noting, "[t]he extent of such fatigue is

III . . . was to insure that 'public' spaces are accessible to the disabled—not to regulate the entirely private areas and activities to which the public does not have access."[142] By suing for a waiver of a rule that applies to every other athlete, Martin is trying to force the private entity to change its competitions into a different kind of competition, one which will fundamentally alter the play and possibly the outcome.

The majority looked to the three sets of rules that generally govern golf. The "Rules of Golf," the "hard card," and the "Notices for Competitors" all set out general guidelines for professional golfers.[143] The Court relied on the fact that the rules never expressly prohibit the use of golf carts during competition. The majority stated, "There is nothing in the Rules of Golf that either forbids the use of carts, or penalizes a player for using a cart."[144]

More than any other point, this finding by the Court demonstrates its unfamiliarity with golf in particular and competitive sports in general, underscoring the need for an expert sports body to properly determine what constitutes a fundamental alteration. Within all sports, and especially at the highest level of competition, there are rules and understandings derived from the nature of play and the essence of competition created, in arguably an arbitrary manner, by a governing body.[145] "And that of course is part of the majesty of sport; it is anything the designers of the rules want it to be, autonomous and, within its realm, sovereign. As such, each variation heralds a new form of competition, a new game."[146] Nowhere in the rules of baseball, for example, does it state that a catcher cannot sit on a stool in front of the umpire at times when his knees are hurting.[147] Nor do the rules of football prohibit a player from riding a golf cart on kickoff returns. Just because an activity is not prohibited by the rule book does not make it permitted. In such gray areas, only an expertise in and

---

impossible to quantify but will vary from golfer to golfer, depending upon factors like weather conditions, temperature, the terrain of the golf course, different golfers' psychological ability to cope with stress, the golfers' age and extent of physical conditioning, and so forth."

142. Brief for the ATP Tour, Inc. and the Ladies Professional Golf Ass'n, Amici Curiae at 10, *Martin* (No. 00-24).

143. *Martin*, 532 U.S. at 666-67.

144. *Id.* at 685.

145. Brief for the ATP Tour, Inc. and the Ladies Professional Golf Ass'n, Amici Curiae at 5, *Martin* (No. 00-24), noting,

> every competition or "game" is, in the end, nothing but a set of manufactured rules, with each rule contributing to, and therefore defining, the nature of the competition. Who can recall, or even know, why bases are ninety feet apart? Who can explain why a tennis player who touches the net during play automatically loses the point? These are simply the rules, and the point, of course, is that every rule of athletic competition is, by definition, fundamental to that game, whatever that game might be.

146. Brief for the ATP Tour, Inc. and the Ladies Professional Golf Ass'n, Amici Curiae at 6, *Martin* (No. 00-24), noting, "Indeed, the rules of athletic competition are simply designed to provide a standard against which to distinguish among competitors."

147. Leverenz, *supra* note 5.

sensitivity to the nuances of the sport empowers the decision-maker to consider whether the play has been fundamentally altered.

The argument that golf carts are used by all public courses throughout the country by the average golfer shows that walking is not a fundamental part of the game is simply not a correct analysis. When the same rule is applied to professionals, it changes the entire landscape. The players are competing for a living and any fundamental change will make the playing field uneven. The nature of golf played by the average fan who is out for a nice stroll with friends on the public course is much different than the game that is played by the game's elite, competing on the finest courses in the world for millions of dollars before countless spectators. This version of golf is played by professionals who have worked all their lives to play the game as a means to make money. Golf is a game that decides who wins based on concrete numbers - the competitor with the lowest score wins. To alter the rules for an individual, in a game that is decided by the slightest of margins, inherently puts other competitors at a disadvantage, constituting a fundamental alteration.

Finally, the holding in *Sandison* applies to the *Martin* case as well. In order to determine the fatigue factor and the way it applies to walking the course during tournaments would also constitute an undue burden on the PGA Tour. In *Sandison*, the Sixth Circuit held it would create an undue burden to require the MHSAA to compare the Plaintiff's effect on competition to each competitor as well as to each competing team.[148] Applying that standard to this case, it would require the PGA Tour to compare the fatigue factor of Martin to each of the other competitors and determine whether allowing Martin to ride a cart would put another competitor at a disadvantage. Would a chronic back problem warrant an exemption from the rule or disadvantage other players by justifying use of a cart? Each one of these cases would require the PGA Tour to evaluate how that disability effects every other competitor. That is simply an undue burden and not required under the ADA.

VI. POSSIBLE IMPLICATIONS ON PROFESSIONAL SPORTS

According to the PGA Tour, all competitors are required to walk during the four-day tournaments. The exhaustion that results from walking is intended to be part of the physical demands of the game. If a golfer does not experience that taxing part of the tournament others are forced to endure, he has gained a physical advantage. This advantage will lead to better shots and lower scores. Every aspect of the game that may not seem important to the recreational golfer may be essential to professional play and can only be determined by those who play at that elite level.

It was clear from the beginning of Casey Martin's case against the PGA Tour in 1998 that other professional sports organizations were concerned that a ruling in favor of Martin would be detrimental to their organizations as well.[149] What

148. 64 F.3d at 1035.

149. Brief for the ATP Tour, Inc. and the Ladies Professional Golf Ass'n, Amici Curiae at 1,

impact the Martin decision will have on the judicial landscape and as a precedent for allowing the judiciary to control fundamental rules of a sport is still largely unclear.[150] Although persistence and overcoming adversity is admirable and everyone deserves the opportunity to play golf, playing golf on a professional tour is a privilege only an elite group can achieve.

"The true challenge courts face in addressing the ADA in competitive situations is . . . attempting to balance the fundamental notion of fair play in competition with the ADA's fundamental goal of full participation."[151] There is a fine line the courts must not cross. "However laudable it might be for courts to create new sporting competitions, with constantly adapting rules so that disabled individuals might be able to compete more effectively, doing so is clearly beyond the intended scope of the ADA."[152] While trying to achieve the goals of the ADA, they should not alter fundamental aspects of a sport and not put other competitors at an inherent disadvantage by trying to give more than *equal access* to an individual.

To prevent future cases where the ADA places an individual at a competitive advantage over others, all parties have new responsibilities as a result of this decision. First, the judicial system should differentiate the application of the reasonable accommodation requirement between the traditional employment sphere and competitive athletic settings.[153] Secondly, organizations must make their rules clear and consistent at all levels of play.[154] As shown in this case, courts may interpret different rules for different levels of play as non-essential to the sport. Although these strategies may not seem like the best way of promoting athletics, it appears this is the only way for a governing authority to protect itself from judicial interference.[155]

Martin (No. 00-24), noting that

> The ATP Tour and the LPGA are similar to the PGA TOUR as that organization is the governing body in men's professional golf in the United States. The ATP Tour and the LPGA have an interest in how the Americans With Disabilities Act (the "ADA") is applied to professional sports competitions.

150. Parent, *supra* note 20, at 145. "There are indications, however, that it has helped foster a sentiment within society that becoming a professional is a right to those who work hard, rather than a privilege offered to those who possess all of the necessary qualifications."

151. Long, *supra* note 15, at 1377.

152. Brief for the ATP Tour, Inc. and the Ladies Professional Golf Ass'n, Amici Curiae at 11, *Martin* (No. 00-24).

153. Long, *supra* note 15, at 1380. "When one leaves the sphere and enters more competitive settings—settings in which there are clearly defined winners and losers and in which an artificial advantage for one participant necessarily means a disadvantage for others—issues of fundamental fairness become more troubling." *Id.*

154. PGA Tour, Inc. v. Martin, 532 U.S. 661, 705 (2001). "The lesson the PGA Tour and other sports organizations should take from this case is to make sure that the same written rules are set forth for all levels of play, and never voluntarily to grant modifications. The second lesson is to end open tryouts." (Scalia, J., dissenting).

155. *Id.* "I doubt that, in the long run, even disabled athletes will be well served by these

Many anticipate the opening of the "floodgates" as a direct result of this decision by the Supreme Court. Athletes from all sports with various disabilities may attempt to change the rules of competitive settings and thereby change the nature of competitive sports. That should not be a concern as a result of *Martin*. First, the ADA specifically defines "disability" and limits what courts may consider as a disability under the statute.[156] The court looks at "the nature and severity of the impairment, the duration or expected duration of the impairment, and the permanent or long term impact, or the expected permanent or long term impact of or resulting from the impairment" to determine whether a given disability restricts the manner in which a person can participate in a major life activity.[157] Minor disabilities will not be recognized by the ADA, essentially eliminating any risk of a litigation frenzy. After all, the PGA Tour never argued that Martin was not "disabled" under the ADA. The real issue was whether the PGA Tour should be held to the ADA standards as a specialized, private entity.

A recent decision by the Supreme Court, despite falling under Title I of the ADA, may narrow the reach of the ADA and prevent the floodgates of litigation from opening as a result of *Martin*. In January, 2002, the Supreme Court held in *Toyota Motor Manufacturing, Kentucky, Inc. v. Williams*[158] that "the [ADA] only covers impairments that affect a person's daily life and does not apply to conditions that prevent a worker from performing a specific job-related task."[159] Ella Williams stated that she was fired from a Toyota plant in Kentucky because a painful repetitive-stress injury to her arms and hands prevented her from doing her job: sponging 500 cars a day. The Court stated that "[m]erely having an impairment does not make one disabled for purposes of the ADA."[160]

The Supreme Court had to decide what an employee must prove to demonstrate that she is substantially limited in performing manual tasks. The Court found that it is not enough for an employee to show that she cannot perform the manual tasks her job requires. "To be substantially limited in the specific major life activity of performing manual tasks, . . . an individual must have an impairment that prevents or severely restricts the individual from doing activities that are of central importance to most people's lives. The impairment's impact must also be permanent or long-term."[161] The Court pointed out that some jobs require unique manual tasks that are not necessarily an important part of most people's lives and to say that people who cannot perform these tasks are

---

incentives that the Court has created." (Scalia, J., dissenting).

156. Weinberg, *supra* note 34, at 773 (citing 42 U.S.C. § 12102(2) (defining "disability" as "(A) a physical or mental impairment that substantially limits one or more of the major life activities of such individual; (B) a record of such impairment; or (C) being regarded as having such an impairment.")).

157. *Id.* (citing 28 C.F.R. § 36.104(1)).

158. 2002 WL 15402 (U.S. 2002).

159. Edward Walsh, *Supreme Court Narrows Reach of Disability Law*, WASH. POST, Jan. 8, 2002, *available at* http://www.washingtonpost.com/wp-dyn/articles/A13467-2002Jan8.html.

160. Toyota, 2002 WL 15402, at *8.

161. *Id.* at *1.

disabled would expand the reach of the ADA beyond what Congress intended when it passed the law.[162]

Toyota focused on what constitutes a disability under the ADA. *Martin,* on the other hand, never addressed whether Martin was disabled under the statute and the issue was never challenged by the PGA Tour. While the *Martin* decision did not influence *Toyota*, the ruling in *Toyota* will severely impact the potential lawsuits from various athletes requesting a waiver of a rule under the ADA. The initial danger of the floodgates opening to more lawsuits within athletic competitions that many people expected after *Martin* should no longer be a major concern after *Toyota.* Unlike Martin, athletes will first have to prove to the court that the ADA covers their disability before they attempt to argue the issues presented in *Martin.*

Potential implications resulting from *Martin*, therefore, should not be the fear of more golfers wanting permission to ride carts during competition. Instead, the most dangerous implication of the *Martin* decision is the role the judiciary plays in deciding whether a waiver of a sports rule would fundamentally alter the nature of the competition. If nothing more, *Martin* demonstrates how this expansive power of the Court to decide what is fundamental to a sport is beyond the Court's judicial expertise and should be beyond its reach as well.[163] Does a Plaintiff have a disability under the ADA? Is a private entity considered a place of public accommodation under Title III, and, if so, what are its duties under the various provisions of the ADA? These are all questions that only a court can decide. It must, however, give ample deference to a governing body on what constitutes a fundamental alteration in this self-contained universe of sports.[164]

---

162. *Id.*

163. Brief for the ATP Tour, Inc. and the Ladies Professional Golf Ass'n, Amici Curiae at 7, 8, *Martin* (No. 00-24), noting,

> only those charged by tradition or agreement with enforcing the rules of the game may amend them, else the game itself is fundamentally changed. Since the game is but a collection of its rules, it is, in the end, only what those in charge say it is. The "keepers" of the game must therefore have ultimate control over what does or does not affect the nature of the competition, much the same way an umpire or referee is, by tradition and agreement, the final arbiter of an on-field dispute. It is not so much whether the runner stealing a base was "safe at second," or whether the tennis ball hit the line, as whether the umpire thought it so. In other words, the rule-making is itself part of the game, and its impact cannot be measured by objective criteria (or even, in the case of the stolen base or the long forehand, by absolute truth). It is only the convention that all will follow the same rules, determined by the same governing body, that gives meaning to the game.

164. PGA Tour, Inc. v. Martin, 532 U.S. 661, 689 (2001). In asserting that the condition of walking is a substantive rule of competition and that waiving it as to any individual would fundamentally alter the nature of the competition, the PGA Tour's evidence included the testimony of the greatest golfers in history. Arnold Palmer, Jack Nicklaus, and Ken Venturi explained that fatigue can be a critical factor in a tournament, particularly on the last day when psychological pressure is at a maximum. They explained to the Court that allowing one person to use a cart might

The power the Court gave itself in making this decision is a dangerous trend for all professional organizations, which in turn will affect the rules of the sport at all levels of competition.

CONCLUSION

The ADA is an essential statute that, when applied correctly, can provide opportunities to disabled individuals and not put other employees or competitors at an inherent disadvantage. However, as Tiger Woods, the world's greatest golfer, observed, the slightest change can have a large impact in a game where small changes are magnified.[165] Courts must distinguish between recreational applications of the ADA and those associated with professional and competitive settings.

The ADA, for all its good intentions, should have only limited applicability to professional sports. As long as the ADA does not fundamentally alter the nature of the event—a question to be decided in the first instance by the governing body that has the expertise and intimate perspective on the sport—the ADA is needed to provide equal access. It is essential, however, that when applying the ADA to professional sports, courts understand that

> [t]he long consistent history of requiring adherence to uniform rules, both in golf and in other sports, reflects a shared understanding of what high-level athletic competition is all about: a test of physical proficiency for different competitors under identical rules. It follows, therefore, that any waiver of a substantive rule for a given competitor is out of keeping with the fundamental premise of professional sports.[166]

It is when *equal access turns into an equal opportunity to win* that the ADA is misapplied and the spirit of competition is crushed. Without fair and equal rules, the once-level playing field of competition is altered, to the detriment of all players, no matter their ability.

---

give a player an advantage over other players who must walk.

165. WOODS, *supra* note 1.

166. Brief for Petitioner at 34, *Martin* (No. 00-24).

PERMITTED USE OF PATENTED INVENTIONS IN THE UNITED STATES: WHY PRESCRIPTION DRUGS DO NOT MERIT COMPULSORY LICENSING

KIRBY W. LEE*

INTRODUCTION

On October 16, 2001, in the wake of several mail anthrax cases, Attorney General John Ashcroft and U.S. Senator Charles Schumer (D-NY) urged the federal government to sanction the generic manufacture of an antibiotic to combat the disease, despite existing patent rights on the drug.[1] Bayer AG manufactures ciprofloxacin, known commonly as Cipro, and has a patent on the drug. "One [sixty]-day supply of Cipro costs just under $700 in the United States, while a [sixty]-day supply of generic ciprofloxacin—not allowed until 2003 in the United States—costs about $20 in other countries . . . ."[2] Although Bayer pledged to increase production of this antibiotic threefold in response to the recent threats of bioterrorism, some public officials still insisted on more action.[3] Senator Schumer initiated talks with three generic drug manufacturers over the possible expedited approval of generic ciprofloxacin, believing that the federal government does have the power to override Bayer's patent rights.[4] A spokesman for the Department of Health and Human Services, however, hesitated at such dramatic action, stating "[w]e have to be careful about patent protections—there's a balance there."[5]

In August 2001, the Brazilian government announced plans to disregard patent rights granted to the Swiss pharmaceutical company Roche for an AIDS drug.[6] Viracept, the brand name for nelfinavir, is an expensive drug often used in AIDS cocktail treatments. Brazil purportedly spends $88 million annually on Viracept alone, which accounted for over a quarter of the country's AIDS program budget.[7] Under mounting pressure to lower the cost of the drug and

---

* J.D. Candidate, 2003, Indiana University School of Law—Indianapolis; B.S., 1997, Purdue University, West Lafayette, Indiana. The author is a patent agent at Eli Lilly and Company. The views expressed herein are those of the author and not necessarily those of Eli Lilly and Company. The author would like to thank Mr. Robert A. Armitage, Senior Vice President and General Counsel, Eli Lilly and Company, for his invaluable guidance.

1. *See N.Y.'s Schumer Wants Anthrax Drug for Government Use Despite Patent*, CNNMONEY, Oct. 16, 2001, *available at* http://money.cnn.com/2001/10/16/news/generic_cipro/ [hereinafter *Schumer Wants Anthrax Drug*].

2. Julie Rovner & April Fulton, *Senate Panel Approves Bill On Generic Drug Availability*, NAT'L J.'S CONGRESSDAILY, Oct. 18, 2001, *available at* LEXIS, News Group File, http://lexis.com.

3. *See id.*

4. *See Schumer Wants Anthrax Drug*, *supra* note 1.

5. *Id.*

6. *See Brazil Plans to Ignore Patent on AIDS Drug*, THE GLOBE AND MAIL, Aug. 23, 2001, *available at* LEXIS, News File, http://lexis.com.

7. *See* Anthony Faiola, *Brazil to Ignore Patent on AIDS Drug*, WASH. POST, Aug. 23, 2001, at A20.

increased criticism from AIDS activists worldwide, Roche eventually reached an agreement with Brazil to lower the price to roughly thirty percent of the price in the United States.[8] Merck & Company, a U.S. pharmaceutical company, also reduced the price on two of its AIDS drugs, indinavir and efavirenz, by approximately sixty percent in anticipation of similar pressure.[9]

Brazil is not the only country to exhibit such a dismissive attitude toward international protection of intellectual property, particularly with respect to prescription AIDS drugs. Previously, South Africa faced the same situation with its own national AIDS crisis, in which approximately 70,000 HIV-positive children are born each year.[10] Pleas to pharmaceutical companies created a stir among intellectual property authorities and human rights activists alike.[11]

Due at least in part to events in Brazil and South Africa, U.S. Representative Sherrod Brown (D-OH) introduced The Affordable Prescription Drugs and Medical Inventions Act.[12] This legislation seeks to amend existing patent laws and allow compulsory licenses potentially applicable to "any invention relating to health care."[13] That is, the government would permit the use of patented inventions, forcing those patent holders to either proactively negotiate licenses or claim reparations after the permitted use.

Although the United States has avoided similar proposals in the past and public health emergencies have been thought to be solely third-world concerns, the recent anthrax scare has revived compulsory licensing arguments with renewed vigor and urgency. On November 6, 2001, Representative Brown appropriately reintroduced his compulsory licensing proposal under a different title, The Public Health Emergency Medicines Act.[14] Deriving many of its provisions from H.R. 1708, this new bill marks an attempt to capitalize on the threat of bioterrorism and feared public health disaster to usher in a compulsory licensing scheme within the U.S. patent system. Public sentiment regarding the rising costs of health care brings prescription drug prices, pharmaceutical patent rights, and compulsory licensing to the forefront of medical, ethical, and economic debate.

This Note argues that compulsory licensing for prescription drugs under these proposed bills is not warranted. It further discusses the rationales that support compulsory licensing and how they are already addressed by other legislative and judicial means. Part I of this Note provides an overview of United States patent law and the transactional interests of government and inventors in

---

8. Jennifer L. Rich, Business/Financial Desk, *Roche Reaches Accord on Drug with Brazil*, N.Y. TIMES, Sept. 1, 2001.

9. *See id.*

10. Barnaby Phillips, *South Africa Sued Over Aids Drugs*, BBC NEWS, Mar. 15, 2001, *at* http://news.bbc.co.uk/hi/english/world/africa/newsid_1502000/1502518.stm.

11. *See id.*

12. H.R. 1708, 107th Cong. (2001). Representative Brown introduced an earlier version of this bill, entitled the Affordable Prescription Drugs Act, H.R. 2927, 106th Cong. (1999).

13. H.R. 1708.

14. H.R. 3235, 107th Cong. (2001).

a patent system. Part II explores the general arguments regarding compulsory licensing within the United States patent system. Part III examines de facto compulsory licensing in the United States, including existing statutory exceptions such as the Drug Price Competition and Patent Term Restoration Act of 1984[15] and government use of intellectual property. Further, Part III will discuss judicial actions that, under special circumstances, essentially result in a compulsory licensing arrangement. Part IV compares U.S. legislation with current existing international treaties and agreements governing patented inventions, especially provisions therein that allow compulsory licensing. Finally, Part V addresses the motivation behind both The Affordable Prescription Drugs and Medical Inventions Act and the Public Health Emergency Medicines Act, analyzes the arguments for compulsory licensing of pharmaceutical drugs, and discusses why these reasons fail in light of other currently available avenues for the permitted use of patented inventions.

I. TRANSACTIONAL INTERESTS OF GOVERNMENT AND INVENTORS IN UNITED STATES PATENT LAW

United States patent law finds its roots in the Constitution, which empowers Congress "[t]o promote the Progress of Science and useful Arts, by securing for limited Times to Authors and Inventors the exclusive Right to their respective Writings and Discoveries."[16] When the U.S. government issues a patent, it includes "a grant . . . of the right to exclude others from making, using, offering for sale, or selling the invention throughout the United States or importing the invention into the United States . . . ."[17] In order to stimulate technological growth and advances, the government essentially grants a limited period of exclusivity to inventors who bring forth and disclose their work.[18] Government provides this as an incentive for inventors to invest time, resources, and money into the innovation process that is often costly. In theory, the public benefits from the introduction of this new and useful invention; inventors, on the other hand, benefit from a period of exclusivity during which they can seek to recoup their investments and profit from their rights.

A. Contract Theory of Patent Law

Many judges and scholars have regarded the modern U.S. patent systems as a type of contract between government and the inventor.[19] The inventor presents

---

15. Pub. L. No. 98-417, 98 Stat. 1585 (codified at as amended in scattered sections of 15, 21, 28, and 35 U.S.C.).

16. U.S. CONST. art. I, § 8, cl. 8.

17. 35 U.S.C. § 154(a)(1) (2000).

18. *See* Lisa A. Huestis, *Patent and Antitrust Law: The Second Circuit Strives Toward Accommodation,* SCM Corporation v. Xerox Corporation, 48 BROOK. L. REV. 767, 773 (1982).

19. *See* Kewanee Oil Co. v. Bicron Corp., 416 U.S. 470, 489 (1974) (stating that a patent is a social contract or bargain, granting exclusive rights in return for public disclosure); *see also In re* Bayer, 568 F.2d 1357 (C.C.P.A. 1978); *In re* Tenney, 254 F.2d 619, 623 (C.C.P.A. 1958)

to the public something that is useful,[20] novel,[21] and unobvious;[22] this public disclosure is his consideration in the bargain. In return, the government provides consideration of exclusive rights to the claimed invention for a limited time.[23] Under this construction, invoking theories of contract law, courts have found patents invalid on the grounds that inventors did not contribute to the public domain anything that was not already known, thus amounting to a failure of consideration.[24] As the inventor's part of the bargain is unsatisfied, courts effectively revoke the government's consideration and are unwilling to enjoin an allegedly infringing party based on the fatally deficient contract.[25] On the other hand, compulsory licensing has been viewed as a failure of consideration on the part of the government. So, even though the inventor satisfied the bargain by publicly divulging his invention, the government's consideration of patent exclusivity is revoked.[26]

Impliedly, however, a patent holder is under no obligation to make, use, sell, or import this invention. Neither the Constitution nor statutory law explicitly requires that the patentee make use of the invention or ensure that the invention is used to its fullest potential. However, some scholars argue that utilization and practical application of the invention is also part of the patent bargain, so that non-use would be a failure of the inventor's consideration in the patent bargain. They argue that strict enforcement of the patent right to exclude, in cases of non-use, does not truly further the spirit of the Patent Clause in the promotion of the useful arts.

B. Is Non-use a Failure of Consideration?

Early in the Twentieth Century, the U.S. Supreme Court recognized the rights of patentees to exclude others even if the patentee himself was not using the invention.[27] At the time, circuit courts were split over the effect of patent non-use with regard to enforcement of a patentee's exclusive rights.[28] Some circuits insisted that use of the patented invention was an incumbent responsibility of the inventor; if an inventor did not use the invention, other parties were free to practice the invention without threat of an action for patent

---

(stating that an inventor must give to the public something it does not already have in consideration for exclusive patent rights). DONALD S. CHISUM, PATENTS (1997).

20. *See* 35 U.S.C. § 101 (2000).

21. *See id.* § 102.

22. *See id.* § 103.

23. Patents are generally subject to a grant of rights extending for a period of twenty years from first filing a patent application. *See id.* § 154(a)(1)-(2).

24. *See Tenney*, 254 F.2d at 622-24.

25. *Id.*

26. *See id.* at 622-23.

27. Continental Paper Bag Co. v. Eastern Paper Bag Co., 210 U.S. 405 (1908).

28. *See id.* at 425-26.

infringement.[29] In *Continental Paper Bag Co.*,[30] the petitioner alleged that non-use of a patent was sufficient grounds to overcome the inventor's protection of exclusive rights. Writing for the majority, Justice McKenna disagreed with this premise, recognizing instead that a patent holder is under no obligation either to use the invention himself or license the invention to others.[31]

> If he [a patentee] sees fit, he may reserve to himself the exclusive use of the invention or discovery. If he will neither use his device nor permit others to use it, he has but suppressed his own, . . . his title is exclusive, and so clearly within the constitutional provisions in respect to private property that he is neither bound to use his discovery himself or permit others to use it.[32]

The Court explained that the right to exclude others was independent of the patent holder's own use or non-use of the patented subject matter.[33] Furthermore, a patent holder is not obligated to license the invention to other interested parties should he choose not to make, use, sell, or import the invention himself.[34]

The Court did, however, acknowledge that some forms of non-use could be directed at wrongful purposes and that such non-use might merit revocation of exclusionary patent rights.[35] Although the Court did not expressly identify these situations at the time, several of these reasons have since developed as new and evolving technology continues to challenge the foundation of patent rights in U.S. law.

II. ARGUMENTS FOR AND AGAINST COMPULSORY LICENSING IN THE UNITED STATES

This section will focus on two leading arguments supporting compulsory licensing of patented inventions: economic benefit and public interest. Advocates of compulsory licensing highlight the supposed economic advantages of compulsory licensing and the evils of perceived monopolistic and anticompetitive behavior that the patent system encourages. Additionally, with particular respect to prescription drugs, arguments in favor of compulsory licensing generally emphasize moral and ethical concerns, citing such circumstances as the crippling spread of certain diseases, unavailability of critical lifesaving medication, and high yet preventable mortality rates.

---

29. *See id.*

30. 210 U.S. 405 (1908).

31. *Id.* at 427-29.

32. *Id.* at 425 (alteration in original) (quoting Bement v. Nat'l Harrow Co., 186 U.S. 70, 90 (1902)).

33. *Id.* at 429. *See also* Hartford-Empire Co. v. United States, 323 U.S. 386, 432 (1945) (holding that a patent owner is under "no obligation . . . to use [the patented invention] or to grant its use to others.").

34. *Continental Paper Bag Co.*, 210 U.S. at 429. *See also* CHISUM, *supra* note 19.

35. *Id.* at 428-29.

A. Economic Rationale

One aspect of exclusive patent rights that draws criticism is that a patent holder may be incapable of meeting demand for the invention. Although this reality may appear on its face to support the introduction of other suppliers in a market via compulsory licensing, the basis for this argument may be rebutted by exploring the economic impact of granted patent rights.

Allowing a competitor to enter a market destroys the fundamental principle of patent protection: exclusivity to compensate for innovation expenses. Theoretically, a "monopolist reduces output below the level that would be found in a perfectly competitive industry."[36] A patent holder may intentionally undersupply goods to maximize profits. Introducing another competitor into a given market would reduce the patent holder's incentive to undersupply and would thus more fully utilize and commercialize the invention.[37] However, increasing access to patented inventions to the detriment of patentees would undermine the incentive to innovate and would deter research. In fact, the mere possibility of compulsory licensing may reduce the incentive for innovation.[38] In high-risk areas of research and development, brand-name pharmaceutical companies (also referred to as "innovator" companies) often seek to recoup costs associated not only with the invention itself, but also with the many other ideas that require resources and fail. Diminishing the return on such research and development by the threat of compulsory licenses could potentially stifle investment in these areas.[39]

The term "monopoly" is used liberally in patent law to describe the position of a patentee in a given market. However, a more precise definition of the relevant market is necessary to understand a patentee's market power. In one instance, Eli Lilly & Co. (Lilly), a pharmaceutical corporation in Indianapolis, Indiana, had exclusivity over the compound nizatidine (Axid®), a histamine H_2-receptor antagonist useful for treating such gastrointestinal maladies as heartburn or stomach ulcers.[40] Lilly could be thought to have had a theoretical "monopoly" over the relevant market, but this market would be limited to nizatidine. Instead, a more practical analysis of the situation reveals that the relevant market cannot be defined simply as the nizatidine market, but rather as all histamine H_2-receptor antagonists. This includes competitors' drugs such as cimetidine (Tagamet®), famotidine (Pepcid®), and ranitidine (Zantac®). Shrewd adherence to monopolistic practices by any one of these competitors could likely have an adverse effect, dissuading consumers from one product and shifting market share

---

36. P. SAMUELSON & W. NORDHAUS, ECONOMICS 166 (16th ed. 1998).

37. WARD S. BOWMAN, JR., PATENT AND ANTITRUST LAW 3 (1973).

38. *See id.*

39. *See* Kevin Rhodes, Comment, *The Federal Circuit's Patent Nonobvious Standards: Theoretical Perspectives on Recent Doctrinal Changes*, 85 NW. U. L. REV. 1051, 1076 (1991).

40. *See* Axid®: nizatidine capsules, Axid® product label, *available at* http://www.reliantrx.com/pdfs/AxidPI.pdf (last visited Nov. 26, 2001).

to other available alternatives. Overly zealous exercise of a patentee's monopoly position in a competitive industry can actually encourage more aggressive "design-around" efforts by competitors.[41] This is one example in which the supposed "monopolistic" rights of a patent holder translate into a much less powerful economic force when viewed in context of a different "relevant market." Thus, a narrow perspective can easily overestimate the true economic power of a patent.

B. Public Interest

The general premise behind this policy rationale is that patent rights, although important, are not absolute. The essential needs of the society as a whole may outweigh the exclusive rights of an individual patentee. Arguments for overriding patent rights in the public interest typically address matters of public health and welfare.[42] Also, matters of national security and defense are considered to impact the public at large and are often treated similarly.[43] For example, judicial determination of public interest has balanced the health and economic interests of citizens against the exclusive rights of a patent holder.[44]

The arguments allowing use of patented inventions for the public good are admittedly not without merit. As an analogy, an individual's *real* property rights, although generally respected and held in high regard, are not absolute. Throughout history, society has recognized certain situations in which the interests of the many outweigh the rights of the individual. For example, in early Seventeenth Century England, it was acknowledged that the King's intrusion on a citizen's private land to mine saltpeter was permitted.[45] Because the act was for the defense of the all the King's people, the right to enter the private land trumped the individual's interest in property.[46] Likewise, a large urban fire necessitated one city's fire department to destroy an individual's house to spare countless other homes and lives.[47] In that case, the court held that "[a]t such times [of emergency], the individual rights of property give way to the higher

---

41. *See* Rebecca S. Eisenberg, *Proprietary Rights and the Norms of Science in Biotechnology Research*, 97 YALE L.J. 177, 206-07 (1987).

42. *See, e.g.*, Plant Variety Protection Act, 7 U.S.C. § 2404 (2000) (mandating a compulsory license if necessary to ensure an adequate supply of food); Clean Air Act, 42 U.S.C. §§ 7401-7626 (2000) (requiring licensing under reasonable terms of technology to prevent and control air pollution).

43. *See, e.g.*, Atomic Energy Act, 42 U.S.C. § 2183 (2000) (citing the public interest as justification for licensing of patented atomic energy inventions).

44. *See* City of Milwaukee v. Activated Sludge, Inc., 69 F.2d 577 (7th Cir.), *cert. denied*, 293 U.S. 576 (1934).

45. *See* The King's Prerogative in Saltpetre, 12 Cl. 12 (Eng. 1607) (stating that the King's trespass onto private land was privileged and that no compensation was owed to the owner).

46. *Id.*

47. Surocco v. Geary, 3 Cal. 69 (1853).

laws of impending necessity."[48] So, too, have intellectual property rights of the individual in limited circumstances yielded to the benefit of society.

However, an analogous application of a doctrine of necessity to intellectual property rights is not as straightforward. The immediate difficulty with this rationale is the inconsistency in establishing what is in the public interest. The definition of public interest may be subject to change, even within a given country, in light of economic and social values at any given time.[49] Variation in this definition among different courts is common, and among nations, the disparity is even more pronounced.[50] The recent events in the United States involving anthrax and Cipro are an excellent illustration of the susceptibility of the public interest argument to impulsive or irrational reactions to perceived emergencies. Although proponents of compulsory licensing are quick to point out the benefits of such a flexible measure, the more troubling outcome of compulsory licensing is the potential for abuse and manipulation of vague standards.[51] Governments intend compulsory licensing as a means for increasing access to critical patented inventions.[52] However, the consequence of such licensing may be the deterrence of companies to invest.

III. DE FACTO COMPULSORY LICENSING IN THE UNITED STATES

Over the past century, several exceptions have been carved into the exclusivity that patentees enjoyed. In the United States, three broad categories permitting use of patented inventions have emerged: statutory exceptions, sovereign immunity for governmental entities under the Eleventh Amendment, and judicial remedies. This section details the erosion of patent rights and explains how competing interests of government, inventors, and the public are resolved by current U.S. laws. Furthermore, this section argues that these exceptions to exclusive patent rights are based upon factors that are outside the realm of the pharmaceutical industry and therefore are unnecessary for prescription drugs in the current U.S. patent system.

A. Statutory Exceptions

Congress has seized upon certain priorities that serve the public interest, passing legislation that provides for compulsory licensing of patents necessary to further efforts in designated fields of technology. For example, the Atomic Energy Act[53] deals with national defense and security in nuclear materials.[54]

---

48. *Id.* at 73.

49. *See generally* Paul S. Haar, *Revision of the Paris Convention: A Realignment of Private and Public Interests in the International Patent System*, 8 BROOK. J. INT'L L. 77 (1982).

50. *See id.*

51. *See id.*

52. Cole M. Fauver, *Compulsory Patent Licensing in the United States: An Idea Whose Time Has Come*, 8 J. INTL. L. BUS. 666, 671 (1988).

53. 42 U.S.C. § 2183 (2000).

54. This statute states, in pertinent part,

Enacted in 1954,[55] the law reflects the national importance of nuclear power in post-World War II times. A further instance is the Plant Variety Protection Act, which states that a compulsory license is mandatory if necessary to ensure an adequate supply of food.[56] A compulsory sale from farmers of saved seed to other farmers is mandatory.[57] Although Congress intended to preserve compelling societal interests—vital national security matters and humanitarian concerns—the effect and necessity of these statutory compulsory licensing provisions are still questioned today. Indeed, despite these examples, the debate surrounding statutory compulsory licensing is far from settled.

A leading example of the ambiguity of compulsory licensing statutes is the Clean Air Act,[58] passed in 1970 amid an escalating international fuel crisis and a rising trend of ecological awareness. Concerned about pollutants, increasing vehicle emissions, and overall air quality levels, Congress proposed "a national research and development program to achieve the prevention and control of air pollution."[59] Additionally, if technology existed that was vital to an industry to meet the goals of the Act, as determined by government officials, a court order could be sought, "requiring the person who owns such patent to license it on such reasonable terms and conditions as the court, after hearing, may determine . . . ."[60]

At the time, the compulsory licensing provision in the Clean Air Act garnered very little attention.[61] It was believed that Congress had feared that companies could control important pollution control patents and strategically build monopolies by exercising patent rights in view of harsh penalties for violations.[62] Since its inception, however, the compulsory licensing provision has seen little litigation in the courts; arguments concerning its impact have generally been relegated to academia.[63] Unfortunately, lack of resolution in the courts has brought cries of victory from commentators of both sides of the debate. Advocates for compulsory licensing cite the compulsory licensing

---

> [w]henever any patent has been declared affected with the public interest, . . . (1) the Commission is hereby licensed to use the invention or discovery covered by such patent in performing any of its powers under this Act; and (2) any person may apply to the Commission for a nonexclusive patent license to use the invention or discovery covered by such patent . . . .

Id. § 2183(b).

55. Pub. L. No. 703, 68 Stat. 919 (1954).

56. 7 U.S.C. § 2404 (2000).

57. *See id.*

58. 42 U.S.C. §§ 7401-7626 (2000).

59. *Id.* § 7401(b)(2).

60. *Id.* § 7608.

61. *See* Jeffry C. Gerber & Peter W. Kitson, *Compulsory Licensing of Patents Under the Clean Air Act of 1970*, 54 J. PAT. OFF. SOC'Y 650 (1972).

62. *See id.*

63. *See* Kenneth J. Nunnenkamp, *Compulsory Licensing of Critical Patents Under CERCLA?*, J. NAT. RESOURCES & ENVTL. L. 397, 406 (1994).

provision as an example of a provision that ensures that future advances in pollution control are appropriately managed to avoid monopolistic control.[64] They assert that such a clause provides for adequate protection with no apparent adverse effects. On the other hand, critics point out that the immeasurable loss of research and development greatly outweigh any benefits of the licensing provision.[65] It is unclear how such a provision may have deterred investment in pollution control. Also untold are the number of settlements or voluntary licensing arrangements motivated by parties seeking to avoid compulsory licenses, which are usually less favorable to patentees.

In 1996, Congress enacted changes to the patent statutes entitled "Limitation on Patent Infringements Relating to a Medical Practitioner's Performance of a Medical Activity" that limited the enforceability of some medical procedures patents.[66] This statute severely limits the exclusivity of patents claiming medical or surgical procedures. The language of the statute is explicit: "[w]ith respect to a medical practitioner's performance of a medical activity that constitutes an infringement[,] . . . [certain remedial provisions] of this title shall not apply against the medical practitioner or against a related health care entity with respect to such medical activity."[67] In essence, patentees of such procedures are denied any remedy from infringing physicians or hospitals. Remedies that would enjoin practitioners from practicing the invention and enable patentees to recover damages are among those that this statute eliminates.[68] In approving this statutory exception, Congress was especially persuaded by the medical profession's argument that doctors have the ethical and professional duty to share knowledge of new, effective treatments with their patients.[69] Protecting medical procedures through the patent system, proponents argued, would encourage secrecy and inhibit the development of life-saving techniques.[70]

It is critical to note, however, that this statute targets only procedures. Congress expressly exempted pharmaceutical drugs and medical devices from the effects of this Act.[71]

> [T]he term "medical activity" . . . shall not include (i) the use of a patented machine, manufacture, or composition of matter in violation of such patent, (ii) the practice of a patented use of a composition of matter in violation of such patent, or (iii) the practice of a process in violation

---

64. *See* Leroy Whitaker, *Compulsory Licensing—Another Nail in the Coffin*, 2 AM. INTELL. PROP. L. ASS'N Q. J. 155, 163-65 (1974).

65. *See* Nunnenkamp, *supra* note 63, at 406-07.

66. *See* Omnibus Consolidated Appropriations Act of 1997, Pub. L. No. 104-208 § 616 (1996).

67. 35 U.S.C. § 287(c)(1) (2000).

68. *Id.* § 287(c).

69. Cynthia M. Ho, *Patents, Patients, and Public Policy: An Incomplete Intersection at 35 U.S.C. § 287(c)*, 33 U.C. DAVIS L. REV. 601, 606 (2000).

70. *See id.*

71. 35 U.S.C. § 287(c)(2)(A) (2000).

of a biotechnology patent.[72]

Congress was careful to craft this exemption very narrowly around medical and surgical procedures. Patented new chemical entities ("NCEs") and medical devices were outside the intended scope of the amendment. Lawmakers acknowledged that the pharmaceutical sector is unique in its reliance on investment-backed expectations.[73] Medical and surgical procedures are more likely to advance through dissemination to other physicians, hospitals, and universities.[74] On the other hand, the same reasoning does not apply to the exploratory and speculative nature of drug research. The highly competitive and costly industry of drug research and development is one that would not be as productive but for the patent incentive for innovation and investment.[75]

Perhaps the statutory reference most relevant to pharmaceutical drugs, experimental use, rests in the Hatch-Waxman Act.[76] This legislation was the direct congressional response to an infringement lawsuit before the Court of Appeals for the Federal Circuit. In *Roche Products, Inc. v Bolar Pharmaceutical Co.*,[77] a generic manufacturer used the innovator company's approved compound in studies for its version of the drug to seek approval from the Food and Drug Administration ("FDA").[78] The Federal Circuit acknowledged that the purely experimental use of a patented invention, independent of commercial gain, should be exempt from infringement liability.[79] However, despite recognition of this permitted "experimental use," the court narrowly interpreted the statutory provisions that allowed for this type of experimentation.[80] The infringing act was to be independent of activities directed to commercial gain, and should have been limited "for amusement, to satisfy idle curiosity, or for strictly philosophical inquiry."[81] Thus, under strict interpretation of the existing statute, the submission of data to regulatory agencies fell outside the scope of permitted use. The Federal Circuit reversed the district court decision for the generic company, holding that the use of a patented drug by a generic drug company regardless of purpose was an act of infringement.[82] Effectively, "innovator" companies

---

72. *Id.* The term "composition of matter" is recognized as patentable subject matter under 35 U.S.C. § 101 (2000).

73. *See* Bloomberg et al., *Patenting Medical Technology: "To Promote the Progress of Science and Useful Arts,"* 317 NEW ENG. J. MED. 565, 566-67 (Aug. 27, 1987).

74. *See* Wendy W. Yang, Note, *Patent Policy and Medical Procedure Patents: The Case for Statutory Exclusion from Patentability*, 1 B.U. J. SCI. & TECH. L. 5, ¶ 51 (1995).

75. *See id.* ¶¶ 53-54.

76. Pub. L. No. 98-417, 98 Stat. 1585 (codified at as amended in scattered sections of 15, 21, 28, and 35 U.S.C.).

77. 733 F.2d 858 (Fed. Cir.), *cert. denied*, 469 U.S. 856 (1984).

78. *Id.* at 864.

79. *Id.* at 860-61. *See also* CHISUM, *supra* note 19, § 16.

80. *See* 733 F.2d at 864.

81. *Id.* at 863.

82. *Id.*

garnered an extended period of exclusivity because generic manufacturers were forced to wait until after a drug's patent expired before work could start on regulatory approval, a process that could take several years.[83]

Immediately following the *Roche* decision, Congress quickly enacted the Drug Price Competition and Patent Term Restoration Act of 1984, known commonly as the Hatch-Waxman Act.[84] As a compromise between the generic drug industry and innovator pharmaceutical companies, the amendment included provisions that would directly override the Federal Circuit holding. Congress changed the patent infringement laws to permit use of patented inventions "solely for uses reasonably related to the development and submission of information under a Federal law which regulates the manufacture, use, or sale of drugs or veterinary biological products."[85] Generic companies then were allowed to practice patented inventions, including patented drugs, in order to satisfy regulatory submission requirements.

The experimental use exception was based on two general principles intended to further technological advances. It was necessary to work patented subject matter 1) to test the feasibility of another's claimed invention, and 2) to continue to innovate and build upon others' work. In the spirit of "promoting the useful arts," Congress had weighed the innovator companies' interest in protection of drug research investment against the public interest of speeding generic drugs to market.[86] Through the Hatch-Waxman Act, Congress had chipped away at the protection of patented pharmaceutical drugs and lessened the effective period of exclusivity necessary to recoup the cost of years of drug research investment.

B. Federal and State Government Use of Intellectual Property

A second general area of permitted use of patented inventions involves federal or state governmental action. If the federal government infringes a patent, the infringement may amount to a "taking" under the Fifth Amendment, and the patentee is entitled to compensation for the infringing use.[87] Because

---

83. *See* Ralph A. Lewis, Comment, *The Emerging Effects of the Drug Price Competition and Patent Term Restoration Act of 1984*, 8 J. CONTEMP. HEALTH L. & POL'Y 361 (1992).

84. Pub. L. No. 98-417, 98 Stat. 1585 (codified at as amended in scattered sections of 15, 21, 28, and 35 U.S.C.). The Federal Circuit decided *Roche* in May 1994. In response to heavy pressure from the pharmaceutical industry, including both generic and innovator manufacturers, Congress quickly enacted the Hatch-Waxman Act in October 1994.

85. 35 U.S.C. § 271(e)(1) (2001).

86. *See* H.R. REP. NO. 98-857, pt. 2, at 30 (1984), *reprinted in* 1984 U.S.C.C.A.N. 2686, 2714. It is important to note that the Hatch-Waxman Act was the product of much deliberation and compromise between innovator pharmaceutical companies and the generic drug industry.

87. *See generally* Lionel Marks Lavenue, *Patent Infringement Against the United States and Government Contractors Under 28 U.S.C. § 1498 in the United States Court of Federal Claims*, 2 J. INTELL. PROP. L. 389 (1995) (stating that the "takings" clause is an appropriate analysis for government infringement of patented inventions).

patent rights are conferred upon inventors by the United States government, a sovereign nation, these granted rights are subject to the eminent domain of the federal government. As an individual's *real* property rights are subject to eminent domain, so too are *intellectual* property rights in an analogous Fifth Amendment "takings" analysis.[88] In a suit against the federal government for unlicensed use of a patent, a patent holder may recover "reasonable and entire compensation."[89] However, absent from the statute is equitable injunctive relief; injunctions are not available to patent holders against the federal government.[90] The infringing use of patented inventions by the states presents a different problem for patent holders.

Recent Supreme Court cases have directly addressed the issue of sovereign immunity of states. In 1996, the U.S. Supreme Court in *Seminole Tribe of Florida v. Florida*[91] debated the limits on Congress' power to relinquish the sovereign immunity of the states. Invoking the Eleventh Amendment, Congress had attempted to regulate commerce between states and the Indian tribes under the Indian Commerce Clause.[92] The Court determined that Congress did not have the power pursuant to the Commerce Clause to abrogate states' sovereign immunity under its Article I powers. Unless a state consented to suit, it could claim sovereign immunity and avoid liability.[93] The Supreme Court did discuss, however, Congress' power under the Fourteenth Amendment to discharge state sovereign immunity.[94] The Court recognized that the Fourteenth Amendment "expand[ed] federal power at the expense of state autonomy, . . . alter[ing] the balance of state and federal power struck by the Constitution."[95]

The Court revisited this issue within a patent infringement context in *Florida Prepaid Postsecondary Education Expense Board v. College Savings Bank and United States*.[96] In that case, College Savings Bank had patented a method of guaranteeing sufficient funds to cover college tuition costs.[97] It marketed this method in the form of certificates of deposit, named CollegeSure CDS, which were "essentially annuity contracts for financing future college expenses."[98] Florida Prepaid, a state-created entity, imitated the idea and created a comparable system for state universities. College Savings Bank initiated a lawsuit for patent infringement under the Patent and Plant Variety Protection Remedy Clarification Act,[99] and Florida Prepaid moved for dismissal on the grounds of state sovereign

---

88. *Id.* at 393-94.
89. 28 U.S.C. § 1498(a) (2000).
90. *Id.* § 1498.
91. 517 U.S. 44 (1996).
92. *Id.* at 47.
93. *Id.*
94. *Id.* at 59.
95. *Id.*
96. 527 U.S. 627 (2000).
97. *Id.* at 630-31.
98. *Id.* at 630.
99. 35 U.S.C. §§ 271, 296 (1994).

immunity. The Supreme Court reversed the holdings of two lower courts in deciding that Congress had improperly annulled states' sovereign immunity by passing this act. As a rationale, the majority noted that the state use of immunity in federal suits was rare, and that it was similarly uncommon that a state would deprive a patent owner of property without a state remedy.[100]

Since the decision in *Florida Prepaid*, commentators have indicated that the holding will be problematic for patentees, as state jurisdictions remain the only surviving venue for patent infringement suits against state government entities.[101] For example, pharmaceutical research strategists may weigh the high-stakes risks of drug development, and indeed could turn away from universities and research institutions, as these entities derive partial funding from state governments. Not only would the holding of *Florida Prepaid* be applicable to the states themselves, but conceivably the argument could be extended to state actors and other peripheral organizations that derive their authority or funding from state governments. *Florida Prepaid* represents a culmination of High Court decisions that, in light of larger federalism and sovereign immunity ideals, opens the door for state use of patented inventions and further erode the sanctity of patentees' property rights.

C. Judicial Action Resulting in Compulsory Licensing Arrangements

Even if an infringing party cannot find relief in statutory infringement exceptions or within sovereign immunity concerns, the federal court system may craft remedies for the patent holder that result in a compulsory licensing relationship. Although the United States Patent and Trademark Office determines patentability through the examination and prosecution process, validity is not finally decided until a matter is litigated before a federal court.[102] Federal judges have many options in the complex case of a patent infringement lawsuit. A patentee may ask for injunctive relief, that the defendant be enjoined from conducting the infringing acts. In such a prayer for remedy, the courts may consider aspects of equity. On the other hand, the patentee may seek monetary damages for infringement. The statutory provisions for patent infringement remedies are explicit, as permissive language surrounds injunctive relief,[103] while compensatory damages are written with imperative language.[104]

---

100. *Florida Prepaid*, 527 U.S. at 647.

101. *See* Peter S. Menell, *Economic Implications of State Sovereign Immunity From Infringement of Federal Intellectual Property Rights*, 33 LOY. L.A. L. REV. 1399 (2000).

102. *See* Fauver, *supra* note 52, at 667.

103. The statute reads in pertinent part: "The several courts having jurisdiction of cases under this title *may* grant injunctions in accordance with the principles of equity to prevent the violation of any right secured by patent, on such terms as the court deems reasonable." 35 U.S.C. § 283 (2000) (emphasis added).

104. The law providing for compensatory damages is written differently: "Upon finding for the claimant the court *shall* award the claimant damages adequate to compensate for the infringement, but in no event less than a reasonable royalty for the use made of the invention by the

1. Injunctive Relief.—Courts have exercised discretion by withholding injunctive relief in certain cases, even if infringement is found on the part of the defendant. The Federal Circuit has cautioned that injunctive relief is not necessarily granted once infringement is decided.[105] In determining remedies for injured patentees, courts sitting in equity have considered the rationale of economic concerns and public interest, as well as the equitable conduct of the patentee himself.

"A patent owner prevailing on the merits of a patent infringement claim will usually be granted a permanent injunction against future infringement unless the public interest otherwise dictates."[106] Courts have weighed the public interest against interests of the patent holder. In *City of Milwaukee v. Activated Sludge, Inc.*,[107] the courts analyzed the impact of injunctive relief, balancing the health and economic consequences the public would suffer against the protection afforded a patent holder. The patentee in that case sought an injunction to stop the city from further working a patented method and apparatus for sewage purification. The appellate court affirmed the trial court's findings, but refused to allow an injunction. That court considered the severe health risk caused by lack of sewage treatment should an injunction be enforced in reaching its decision.[108] Even today, *City of Milwaukee* represents the leading case in which public interest was found to be compelling in itself to justify denial of injunctive relief.[109]

Courts may also balance the detriment to the infringing party against the benefit to be gained by the patent holder when granting an injunction.[110] Furthermore, a compulsory license may be a possible remedy for an aggrieved plaintiff when the defendant is guilty of antitrust violations.[111] The nature of the

---

infringer, together with interest and costs as fixed by the court." 35 U.S.C. § 284 (2001) (emphasis added).

105. *See* Roche Prods., Inc. v. Bolar Pharm. Co., 733 F.2d 858, 865-67 (Fed. Cir.), *cert. denied*, 469 U.S. 856 (1984).

106. CHISUM, *supra* note 19, § 20 (footnote omitted).

107. 69 F.2d 577 (7th Cir.), *cert. denied*, 293 U.S. 576 (1934).

108. *Id.* at 593.

109. *See also* Vitamin Technologists, Inc. v. Wis. Alumni Research Found., 146 F.2d 941, 945 (9th Cir. 1944) (concluding that the public interest of production of oleomargarine, the "butter of the poor," outweighed the patent holder's interest in retaining exclusive rights); Hybritech, Inc. v. Abbot Labs., 4 U.S.P.Q.2d 1001 (C.D. Cal. 1987) (denying patentee injunctive relief despite infringement because it was in public interest to continue production of infringing medical test kits that patentee was not itself marketing). *Cf.* Wis. Alumni Research Found. v. Gen. Elec. Co., 880 F. Supp. 1266, 1277 (E.D. Wis. 1995) (granting a permanent injunction because of the public interest in preventing infringement of valid patents). Courts have, however, construed the term "public interest" to include the guarantee of certainty and enforceability to patent holders.

110. *See* Am. Safety Device Co. v. Kurland Chem. Co., 68 F.2d 734 (2d Cir. 1934).

111. *See* United States v. U.S. Gypsum Co., 340 U.S. 76 (1950); United States v. Hartford-Empire Co., 46 F. Supp. 541 (N.D. Ohio 1942) (granting compulsory licenses as a remedy for antitrust violations); *see also* Carlisle M. Moore, *A Study of Compulsory Licensing and Dedication*

claim of injunctive relief allows a court flexibility in deciding the most appropriate sanction in a successful infringement lawsuit. This equitable determination is a suitable measure for permitting use of patented inventions while avoiding the overly broad and generalized reach of a compulsory licensing statute.

2. Damages in a Patent Infringement Action.—In a patent infringement action, a plaintiff may elect to seek damages. Statutory provisions require "in no event less than a reasonable royalty" for infringement.[112] Beyond reasonable royalties, however, a patentee may seek lost profit damages for infringement.

To recover lost profits, "a patent owner must prove a causal relation between the infringement and its loss of profits." The Federal Circuit stated that

> a patentee receives a reasonable royalty for any of the infringer's sales not included in the lost profit calculation. Thus, a patentee may obtain lost profit damages for that portion of the infringer's sales for which the patentee can demonstrate "but for" causation and reasonable royalties for any remaining infringing.[113]

Moreover, during the damages stage of a patent infringement action, a judicial determination of de minimis infringement damages may further limit the relief to which a patentee is entitled. In *Embrex, Inc. v. Service Engineering Corp.*,[114] the defendant was accused of infringing a patented process for injecting a vaccine into an avian egg. After affirming the trial court's finding of infringement, the Federal Circuit vacated the awarded damages of $500,000. "Because the only cognizable infringement in this case [was] the testing and those tests were not shown to cause any loss of profits to Embrex,"[115] the Federal Circuit remanded the case to the trial court for a finding of reasonable royalties.[116] Therefore, in cases concerning mere testing, a patentee may find it difficult to establish sufficient evidence to compute a reasonable royalty. This holding furthers the patent system goal of promoting scientific inquiry by protecting and encouraging research.

IV. THE TRIPS AGREEMENT AND ITS COMPULSORY LICENSING PROVISIONS

Representative Brown's compulsory licensing bills include reference to the Agreement on Trade-Related Aspects of Intellectual Property Rights ("TRIPs").[117] This international treaty of the World Trade Organization

---

of Patents as Relief Measures in Antitrust Cases, 24 GEO. WASH. L. REV. 223 (1955).

112. 35 U.S.C. § 284 (2000).

113. Crystal Semiconductor Corp. v. Tritech Microelectronics Int'l, Inc., 246 F.3d 1336, 1353-54 (Fed. Cir. 1996) (quoting BIC Leisure Prods., Inc. v. Windsurfing Int'l, Inc., 1 F.3d 1214, 1218 (Fed. Cir. 1993)).

114. 216 F.3d 1343 (Fed. Cir. 2000).

115. *Id.* at 1350.

116. *Id.*

117. AGREEMENT ON TRADE-RELATED ASPECTS OF INTELLECTUAL PROPERTY RIGHTS

("WTO") promotes uniformity among member nations by introducing standards for patent protection worldwide. This section details the compulsory licensing provisions of TRIPs and how the proposed legislation is unnecessarily redundant.

The TRIPs agreement does provide for compulsory licensing of patented inventions. The criteria for such compulsory licensing circumstances exist in Article 31 of the TRIPs document. Most notably, Article 31(b) allows compulsory licensing of patented inventions in situations of national emergency or extreme urgency.[118] Article 31(g) provides that use of the patented invention under the license may continue only so long as the original need exists.[119]

There exists an ongoing dispute between developed countries possessing key patented technology and those bearing "developing nation" status that typically claim the greatest need. Developed nations such as the United States generally possess advanced medical technology and resources, and they advocate a narrow interpretation of the TRIPs compulsory licensing provisions. Criticism of TRIPs compulsory licensing provisions centers on the ambiguity and latitude in interpretation. Terms such as "circumstances" and "purpose" could lead to inconsistent application.[120] Nations still may exercise sovereign power by declaring "national emergency." There are few guidelines that indicate standards for such events, and this section of the TRIPs agreement has not been challenged to an authoritative body.[121] Developing nations, however, argue for an expansive reading of Article 31, and present humanitarian issues such as AIDS crises and other public health concerns as justification. Pharmaceutical companies are placed in the awkward position: on one hand, they want to avoid arguing that widespread diseases are not a matter of public interest, but on the other, they are wary of importation or foreign infringement that would result from enactment of compulsory licensing provisions.[122]

Article 31(c) of TRIPs limits licensing of patented inventions to the original purpose for which the license was granted.[123] This condition within the treaty addresses the concern of developed nations that appropriation of patented inventions may be abused beyond the national emergency or circumstances that created the justification for a compulsory license. Specifically, the concern is that even after the emergency need is met, rogue companies will inundate the international market illegally.

Clarification from the WTO regarding the terminology of TRIPs and the boundaries of the compulsory licensing provisions is needed to provide a

---

[hereinafter TRIPs], *available at* http://www.wto.org/english/docs_e/legal_e/final_e.htm (last visited Jan. 3, 2003).

118. *Id.* Part II, sec. 5, art. 31(b).

119. *Id.* art. 31(g).

120. *See* 145 CONG. REC. H6027 (daily ed. July 21, 1999).

121. *See* Robert J. Gutowski, Comment, *The Marriage of Intellectual Property and International Trade in the TRIPs Agreement: Strange Bedfellows or a Match Made in Heaven?*, 47 BUFF. L. REV. 713, 720-24 (1999).

122. *See* 145 CONG. REC. H6027 (daily ed. July 21, 1999).

123. TRIPs, *supra* note 117, art. 31(c).

meaningful international agreement. For example, in the case of patented AIDS pharmaceutical drug therapies, the United States and South Africa argued over the precise application of the TRIPs agreement.[124] However, despite the weaknesses of TRIPs, it does present background for analysis of domestic compulsory licensing laws. The following section analyzes the current legislative proposals before the U.S. House of Representatives.

V. PROPOSED COMPULSORY LICENSING OF HEALTH CARE INVENTIONS

The proposed legislation sponsored by Representative Sherrod Brown seeks to "use market competition to bring down the cost of prescription drugs."[125] Supporters of these bills eagerly cite the success of compulsory licensing provisions in the Clean Air Act.[126] They insist that the rising costs of health care may be curbed by licensing measures for expensive prescription drugs, prices of which "bear[] no resemblance to pricing norms for other industries."[127] The broad sweeping language of this proposed legislation relates to "any invention related to health care,"[128] which encompasses any drug or device, any biological product, or any technology or process to the extent the technology or process is applied to health or health care.[129] It is unclear, however, whether Representative Brown's bills strike the proper balance between public access to drug inventions and research incentive.

A. H.R. 1708: The Affordable Prescription Drugs and Medical Inventions Act

This bill bestows to both the Secretary of Health and Human Services and the Federal Trade Commission "the right to establish other use of the subject matter of the patent without authorization of the right holder"[130] for any invention

---

124. *See* Office of the United States Trade Representative, Executive Office of the President, *United States-South Africa Understanding on Intellectual Property*, *available at* http://www.ustr.gov/releases/1999/09/99-76.html (last visited October 15, 2001). Interestingly, both nations agreed to resolve the dispute over intellectual property rights privately as opposed to seeking adjudication from the WTO's Dispute Settlement Board, perhaps for concern of an unfavorable interpretation of the treaty provisions.

125. U.S. Representative Sherrod Brown (D-OH), The Affordable Prescription Drugs Act, Bill Summary, *available at* http://www.house.gov/sherrodbrown/rxdrugsumm.htm (last visited October 15, 2001).

126. Representative Brown refers to the precedent established in the Clean Air Act, discussed *supra*, which provides for compulsory licensing of patented pollution control devices deemed necessary by government to the success of the Act. *Medicare Prescription Drugs: Hearings Before the House Subcomm. on Health and the Environment*, 105th Cong. (1999) [hereinafter *Hearings*] (statement of Rep. Brown, Member, House Comm. on Energy and Commerce), *available at* http://www.house.gov/sherrodbrown/ medpresdrg.htm (last visited Oct. 15, 2001).

127. *Hearings*, *supra* note 126.

128. H.R. 1708, § 2, 107th Cong. (2001).

129. *Id.*

130. *Id.*

related to health care. In order for the government agencies to invoke these licensing rights, such invention must fulfill at least one of five determinative factors.[131] The bill further provides for "adequate remuneration for the use of the patent,"[132] and claims consistency with existing international treaty provisions.[133]

The determinative factors in the proposed bill are directly analogous to existing theories supporting compulsory licensing. The first item relates to the argument that non-use of a patented invention may be grounds for compulsorily licensing the invention. If "[t]he patent holder . . . has not taken, or is not expected to take within a reasonable time, effective steps to achieve practical application of the subject invention in a field of use,"[134] then the patented invention may be subject to a compulsory license under this bill. The proposition that licensing is mandated should the patent holder fail to use the patented invention himself is the very concept dismissed by the U.S. Supreme Court in *Continental Paper Bag Co.*[135] Non-use of patented inventions has not been upheld as a valid justification for compulsorily licensing subject matter of any kind, and the rationale should fail when targeting health care inventions specifically. Moreover, ambiguous terms such as "reasonable time," "effective steps," and "practical application" are subject to a wide range of judicial interpretations that could lead to deterring inconsistencies in enforcement.

Two additional factors invoke the public interest or public health argument. A compulsory license option may be triggered if "[t]he invention claimed in the patent is needed for research purposes that would benefit the public health, and is not licensed on reasonable terms and conditions," or if "use of the subject matter of the patent is necessary to alleviate health or safety needs which are not adequately satisfied."[136] This rationale could be extended to reach many different types of technologies so long as a tie to public health could be established. Furthermore, the impetus of "research purposes" was addressed directly in the experimental use provisions of the Hatch-Waxman Act.[137] Recent decisions from the Court of Appeals for the Federal Circuit have indicated that this argument is disfavored, deferring instead to respect of the patentee's intellectual property rights. The limitation of reasonable licensing terms is also questionable. It is conceivable that new and unobvious innovations in a particular field may indeed merit terms favorable for the patentee; this is the nature of pioneer inventions. By regulating the terms by which parties seek licenses, Congress may very well inhibit the incentive to invest in an industry as costly and research-intensive as pharmaceuticals.

Another factor reflects the judicial denial of equitable relief in cases of

---

131. *Id.*

132. *Id.*

133. *Id.* The bill refers to the TRIPs, discussed *supra*, and the Uruguay Round Agreements Act, § 101(d)(15).

134. H.R. 1708, § 2.

135. 210 U.S. 405, 429 (1908).

136. H.R. 1708, § 2, 107th Cong. (2001).

137. *See supra* Part III.A.

anticompetitive behavior. Traditionally, the United States has frowned upon antitrust-like behaviors, a paradigm often forced into conflict by the exclusionary nature of the patent system.[138] The bill permits compulsory licensing in the event that "the patented invention is priced excessively relative to the median price for developed countries or by other reasonable standards, and that such pricing contravenes the public interest."[139] The United States is responsible for a great majority of the total costs for drug research and development. This portion of the proposed bill aims to target the perceived unfairness in pricing relative to other developed nations. However, advocates for the innovator pharmaceutical companies point out that the United States often reaps the benefit of life-saving therapies years ahead of other nations. They defend discrepancies in drug pricing compared to other industrialized nations by citing the advances of the U.S. health care system and the higher standard of living enjoyed by the average U.S. citizen. In fact, this type of pricing comparison would be difficult to weigh practically and even more difficult to implement.

The final factor permits compulsory licensing if "[a]n invention covered by a [second] patent . . . cannot be exploited without infringing upon the [first] patent . . . , insofar as the invention claimed in the second patent involves an important technical advance."[140] This factor relates to the patent misuse doctrine, a common law principle raised during litigation. Patent misuse is available as an affirmative defense in a patent infringement action, as alleged patent infringers assert that the plaintiff-patentee has abused the patent grant. The allegation is that the patentee has overreached and attempted to extend its exclusivity to items that are not within the scope of the patent.[141] If successful, the affirmative defense can result in the denial of equitable relief.[142]

The determinative factors cited in The Affordable Prescription Drugs and Medical Inventions Act are wholly redundant and unnecessary. It may be argued that these considerations merely represent the codification of common law principles. However, application of these remedies under the aforementioned circumstances is by no means universal or automatic in patent infringement cases, and health care inventions do not merit special consideration of this option. H.R. 1708 presents a backwards step for pharmaceutical innovation and public health concerns.

---

138. *See* Philip Girard, *Impact of United States Antitrust Laws on Territorially-Limited International Patent Licensing Agreements*, 11 U. S.F. L. REV. 640 (1977). Judges have, on occasion, considered the anticompetitive behavior of the patent holder in determining appropriate sanctions for the infringing party. *See* United States v. U.S. Gypsum Co., 340 U.S. 76 (1950); United States v. Hartford-Empire Co., 46 F. Supp. 541 (N.D. Ohio 1942) (granting compulsory licenses as a remedy for antitrust violations).

139. H.R. 1708, § 2.

140. *Id.*

141. *See* CHISUM, *supra* note 19, § 19.

142. *See* Morton Salt Co. v. G.S. Suppiger Co., 314 U.S. 488 (1942). The Patent Act of 1952 revised the statutory law to limit the patent misuse doctrine to tying agreements involving staple products. *See* Dawson Chem. Co. v. Rohm & Haas Co., 448 U.S. 176, 209 (1980).

B. H.R. 3235: The Public Health Emergency Medicines Act— An Even Broader Approach to Compulsory Licensing

Less than two months following the tragic and shocking events of September 11, 2001, Representative Sherrod Brown introduced H.R. 3235, a statutory measure directed to the threat of bioterrorism.[143] This revised legislation also aspires to establish compulsory licensing of patented inventions, but in a much broader sense than H.R. 1708. The proposed statute reads:

> In the case of any invention relating to health care[,] the Secretary of Health and Human Services shall have the right to authorize use of the subject matter of the patent without authorization of the patent holder or any licensees of the patent holder if the Secretary makes the determination that the invention is needed to address a public health emergency.[144]

Absent from H.R. 3235 are the determinative factors of H.R. 1708, which provided at minimum some measure of guidance for a reasonable assessment of applicability. In the case of the Public Health Emergency Medicines Act, however, *any* invention related to health care would be implicated provided that a public health emergency exists. It is not difficult to conceive of the multitude of patented inventions this includes, making this overly broad proposal unrealistic and infeasible.

Representative Brown remarked that his bill "would address the compensation issue [of use of patented inventions], precluding endless court battles and unnecessary government spending."[145] He cited "[u]nencumbered access to drugs [as] an essential element in [the] response to bioterrorism."[146] Yet, he also conceded that "[t]he links between antibiotic resistance and bioterrorism are clear. . . . We can only assume that anthrax, and other bacterial agents, could also be engineered to resist antibiotics—including drugs like Cipro."[147] Under a compulsory licensing scheme as proposed by Representative Brown, the incentives under the current U.S. patent system are severely weakened so that the next generation of drug therapies may never arrive.

CONCLUSION

Compulsory licensing of patented inventions is not merited for pharmaceutical drugs. Proposed bills such as the Affordable Prescription Drugs and Medical Inventions Act and the Public Health Emergency Medicines Act do

---

143. *See* H.R. 3235, § 2, 107th Cong. (2001).

144. *Id.*

145. U.S. Representative Sherrod Brown (D-OH), Remarks on the Public Health Emergency Medicines Act, *available at* http://www.house.gov/sherrodbrown/bioterror1115.htm (last visited Feb. 25, 2002).

146. *Id.*

147. *Id.*

not take into account the present range of legislative and judicial avenues for relief that are available. Existing remedies already satisfy arguments concerning the public interest and economic reasons. These arguments are too easily influenced by contemporary sentiment. The recent events in the United States involving Cipro and the threat of anthrax present a prime example of this phenomenon. Proponents of compulsory licensing are too quick to point to perceived health emergencies and urgent needs while ignoring the deterrence on innovation and the continued erosion of patent rights. In past legislation, Congress has correctly recognized the unique incentive-backed investment expectations of the pharmaceutical industry and should wisely avoid these broad, sweeping compulsory licensing bills. Without the preservation of exclusionary patent rights for pharmaceuticals, there may not be a next generation of critical drugs to meet future needs.

"DUEL" BANKING SYSTEM? STATE BANK PARITY LAWS: AN EXAMINATION OF REGULATORY PRACTICE, CONSTITUTIONAL ISSUES, AND PHILOSOPHICAL QUESTIONS

JOHN J. SCHROEDER*

INTRODUCTION

Depository financial institutions in the United States, including banks, credit unions, and thrifts, are unique in that their incorporators and/or management have a choice between state and federal charters, regulatory authorities, and governing statutes. No other industry has separate and distinct laws governing its powers, regulation, and organizational structure. This phenomenon is known as the "dual banking system."[1] Every state has an agency, or agencies, that charter and regulate these three types of financial services providers.[2] Alternatively, federal charters for banks, thrifts, and credit unions are provided by the Office of the Comptroller of the Currency ("OCC"), the Office of Thrift Supervision ("OTS"), and the National Credit Union Administration ("NCUA"), respectively.[3] For reasons that will be discussed below, the availability of this choice of charters has contributed greatly to the industry innovations and the expansion of powers that financial institutions have experienced in this country. It has also provided necessary "checks and balances," ensuring against oppressive regulation. Further, the system fosters a competitive environment between state and federal regulators. This healthy competition and the "level playing field"[4] it fosters are essential to the survival of the dual banking system.

State bank parity laws have been one means by which states have striven to provide a charter choice that meets the needs of its regulated banks, is

---

* J.D. Candidate, 2003, Indiana University School of Law—Indianapolis; B.S., 1986, Indiana University Kelley School of Business, Indianapolis, Indiana. The author is the Supervisor of the Administration Division for the Indiana Department of Financial Institutions. The views and opinions of the author are his own, and do not necessarily reflect those of the Indiana Department of Financial Institutions.

1. *See generally* Arthur E. Wilmarth, Jr., The Dual Banking System—A Legal History (Sept. 30, 1991) (unpublished paper presented at the Education Foundation of State Bank Supervisors (EFSBS) Seminar for State Banking Department Attorneys, on file with author).

2. Christian A. Johnson, *Wild Card Statutes, Parity, and National Banks—The Renascence of State Banking Powers*, 26 LOY. U. CHI. L.J. 351, 357 (1995); Raquel Maria Prieguez, *Federal Common Law and the Need for Uniformity in the Regulation of Federal Savings and Loan Associations and Federal Credit Unions*, 29 SAN DIEGO L. REV. 743, 751, 755 (1992).

3. Jerry W. Markham, *Banking Regulation: Its History and Future*, 4 N.C. BANKING INST. 221, 228 (2000); Prieguez, *supra* note 2, at 749, 786-88.

4. Kenneth F. Ehrlich, *Gramm-Leach-Bliley: Federal Preemption of Massachusetts Bank Insurance Sales Rules?*, 20 ANN. REV. BANKING L. 121, 125 (2001) (describing the intention of 1997 federal legislation to foster competition and equality between national and state charters). *Id.*

competitive with the federal alternative, and promotes "safety and soundness"[5] in the industry. Parity laws provide state regulators and lawmakers a flexible and timely method of expanding and/or amending the permissible powers of state-chartered banks in response to newly adopted federal initiatives.[6] This adaptability is particularly important given the fact that the legislatures in many states are in session only part-time.[7] The ability for states to adapt has become increasingly important in recent years, as federal regulators have aggressively interpreted their authority to expand the powers of federally chartered financial institutions.[8] While many of the issues to be addressed in this Note apply equally to all three types of traditional depository institutions—banks, thrifts, and credit unions—in order to keep the topic manageable, I will concentrate specifically on the bank charter.

Part I of this Note provides a brief history of the "dual banking system" in the United States. Included will be a discussion of the positive effects this system has had on the country's banking industry. Part II consists of an analysis of existing state bank parity laws and the various means of their application throughout the fifty states.[9] Included in this section is a discussion of the extent to which the parity laws preempt or simply supplement other state laws, and whether the powers are afforded automatically, or are subject either to the

---

5. *See* Ralph E. Sharpe, *Prompt Regulatory Action and Safety and Soundness Tripwires Under FDICIA*, 625 PRACTICING LAW INST.—COMMERCIAL LAW 217, 236-44 (1992). "Safety and soundness" is the general standard under which bank regulators review the operations of banking companies. This article describes the standards and criteria used in determining banks' conformity to safe and sound practice.

6. *See* Letter from James B. Kauffman, Jr., Acting Secretary of Banking, State of Pennsylvania, to all Pennsylvania State-Chartered Banks, Banks and Trust Companies, Savings Banks, and Trust Companies (Nov. 29, 2000) (announcing the adoption of their parity provision, opining that it would ensure "a level playing field on which Pennsylvania State-Charters can continue to successfully compete") (on file with author); *see also* Press Release, New York Governor George E. Pataki (July 23, 1998) (recognizing a growing "competitive imbalance" and hailing New York's parity law as ensuring "the State banking charter will remain attractive and competitive") (on file with author).

7. John Devlin, *Toward a State Constitutional Analysis of Allocation of Powers: Legislators and Legislative Appointees Performing Administrative Functions*, 66 TEMP. L. REV. 1205, 1229 (1993). Though the frequency and length of state legislative sessions vary, their part-time nature can leave significant lapses of time between enactments of federal law and consideration by state lawmakers.

8. *See* Arthur E. Wilmarth, Jr., Recent Developments Related to the Preemption of State Laws by the Office of the Comptroller of the Currency ("OCC") and the Office of Thrift Supervision ("OTS") (July 31, 2001) (unpublished paper presented at the Conference of State Bank Supervisors ("CSBS") Legal Seminar) (on file with author).

9. Information compiled from a survey of state banking department representatives, telephone interviews, e-mail correspondence, and independent research, represented in Table I—Summary of State Bank Parity Laws, included *infra* [hereinafter Survey Results] (on file with author).

discretion of the state regulator, or to specific legislative constraints. Following is a discussion of the various types of powers and authorities that have been requested under parity provisions. Included in Part II is an analysis of the interplay between the parity provisions and the "incidental and proper" clauses that are also common in state banking codes. These "incidental and proper" clauses, which sometimes require regulatory approval, provide banks the ability to exercise powers that are not enumerated but are deemed "incidental and proper" to banking. When regulatory agencies interpret these "incidental and proper" clauses broadly, they can serve to expand permissible bank powers even in the absence of parity provisions.[10] Although parity provisions are designed to promote a "level playing field," particularly between state and national charters, "incidental and proper" clauses can be used to seek powers that, while arguably incidental to the business of banking, are not, as yet, available to national banks.

Part III of this Note discusses parity laws from a constitutional perspective. This question first arises in the form of a potential abdication or delegation of lawmaking authority by state legislatures when they provide for the "automatic" extension of theretofore-unauthorized powers to state banks at the discretion of federal lawmakers, or arguably worse yet, federal regulators.[11] This concern is heightened even further when the power that is extended to state banks was not previously simply unauthorized, but specifically prohibited by state law. A second question of delegation involves the constitutionality of a statute that allows a state executive branch agency the latitude to unilaterally expand a theretofore legislatively enacted list of permissible bank powers.

These issues lead to a discussion of philosophical issues in Part IV. Banks, while not public entities, certainly raise significant public policy concerns and benefit from public support, i.e., federal deposit insurance. They are in the business of accepting citizens' money in the form of deposits and investing it in, for example, loans, securities, and real estate. For these reasons, states have always had a strong interest in the powers and activities afforded to banks. Banks historically could only engage in specifically enumerated powers.[12] For better or worse, parity provisions can significantly alter this regulatory structure by expanding these powers beyond those adopted legislatively. In the case of parity provisions that automatically allow a state bank to engage in an activity

---

10. The wording of these "incidental and proper" clauses varies from state to state. Examples include, "[d]o any business and exercise any powers incident to the business of banks," ALA. CODE § 5-5A-18(12) (1990 & Supp. 2001); "exercise all powers incidental and proper . . . in carrying on a general banking business," IND. CODE § 28-1-11-3.1(a) (1998 & Supp. 2001); "all powers incidental to the conduct of banking business," PA. STAT. ANN. tit. 7, § 315(I) (1995 & Supp. 2001).

11. As noted *supra*, note 3, the OCC is the chartering and regulatory authority for national banks. In that role, it is charged with interpreting the National Banking Act, and thus determining permissible powers for national banks. 12 U.S.C. § 24 (2000).

12. *See* Karol K. Sparks, *Banking and Insurance: One Year After Gramm-Leach-Bliley*, SF57 A.L.I.-A.B.A. 667, 671 (2001) (discussing the effect this new law will have on available bank powers).

that was previously unauthorized, or even specifically prohibited, the question becomes, "Why is it a safe business practice now?" This inquiry brings the whole historical practice of specifically enumerating bank powers into question.

A second philosophical issue is the importance of consistency, or lack thereof, among the various states in the adoption and application of parity provisions. In this age of interstate banking, seamless regulation is viewed favorably by large banking organizations. However, any lockstep effort by states could also be viewed as an endorsement of a national regulatory environment and an undermining of the dual banking system.

Parity provisions have played an important role in both the evolution of bank powers and the continued viability of the dual banking system. Their near-unanimous adoption throughout the country is evidence of their importance. Given the consistent state interest in protecting the safety and soundness of financial institutions, particularly in an interstate environment, an effective argument can be made for a more universal application of parity laws across state lines. Further, while the constitutionality of parity provisions can be debated, there have been no significant challenges to them. This is not likely to change given their utility and widespread acceptance.

I. DUAL BANKING SYSTEM HISTORY

A. Structure of Bank Regulation in the United States

A dual banking system has existed in this country since the enactment of the National Banking Act in 1863.[13] Prior to this time, other than the First and Second Banks of the United States, only state banks existed. This structure had spawned several hundred state banks, each issuing their own currency. National banks were authorized in 1863 primarily due to the need to establish a uniform currency to fund the Civil War.[14] The OCC serves as the primary regulator and chartering authority for national banks, and the executive branch of each state maintains an agency charged with chartering and regulating state banks.

In addition, all national banks and virtually all state banks are insured by the Federal Deposit Insurance Corporation ("FDIC"), resulting in an additional regulator for most state banks, and in some instances, for national banks.[15] Further, all national banks and many state banks are members of the Federal

---

13. Markham, *supra* note 3, at 228 (referencing the National Banking Act at 12 U.S.C. § 24 (2000)).

14. THOMAS MAYER ET AL., MONEY, BANKING, AND THE ECONOMY 35-42 (1981).

15. *See* Johnson, *supra* note 2, at 358-61; *see also* Press Release, FDIC Chairman Donald Powell, Statement on FDIC Board Approval of Special Examination Activities (Jan. 29, 2002) (introducing an interagency regulatory agreement entitled "Coordination of Expanded Supervisory Information Sharing and Special Examinations") (on file with author). The agreement was negotiated between the FDIC, the OCC, the Federal Reserve Board, and the OTS, and it expands the circumstances under which the FDIC will conduct examinations of banks not directly supervised by the FDIC.

Reserve System that can result in additional regulation.[16] Also, all banking companies that have adopted a bank holding company structure are subject to regulation by the Federal Reserve Board ("FRB").[17] While the regulatory presence of the FDIC and FRB are not considered a part of the dual banking phenomena, the existence of this multitude of regulators, together with their respective regulations, can complicate the regulatory process for both bankers and regulators.

B. Effects of the Dual Banking System

The dual banking system provides a charter choice for bank management to exercise based on available powers, geographic concerns, accessibility of regulators, regulatory philosophy, and costs. Generally speaking, the larger interstate or international companies have tended to hold national charters. Smaller, community bankers often choose to operate under the more local regulatory environment provided by the state regulator. While these characterizations are only generalities, the numbers tend to support them. As of December 31, 2001, there were 8080 commercial banks in the United States.[18] Of these, 2137 were national banks and 5943 were state banks.[19] The average size of the national banks was $1.7 billion, while the average-sized state bank held $494 million in assets.[20]

Historically, the existence of the dual system has provided for innovation in products and services in the industry. The competitive nature of the dual banking system has prompted individual states to be responsive to the needs of their constituent bankers, thereby resulting in new products and powers. When these responsive innovations are multiplied by the fifty state chartering authorities, the result actually belies the "dual" banking system name and creates numerous opportunities for experimentation. Among innovations attributed to the state system are checking accounts, branching, real estate lending, deposit insurance, and trust services.[21] The OCC has also been responsive, increasingly so in recent years, in authorizing additional national bank powers. National banks, through OCC authorization, have introduced or expanded powers in the areas of insurance brokerage, travel agencies, operating subsidiaries, leasing, and data processing services.[22] More recently, OCC interpretations have provided expanded geographical opportunities for national banks (branching powers), as

---

16. Johnson, *supra* note 2, at 359.

17. *Id.* at 358-61.

18. Federal Deposit Insurance Corporation, FDIC—Statistics on Depository Institutions Report, Assets and Liabilities, *at* http://www2.fdic.gov/sdi/rpt_Financial.asp (last visited May 15, 2002).

19. *Id.*

20. *Id.*

21. Arthur E. Wilmarth, Jr., *The Expansion of State Bank Powers, the Federal Response, and the Case for Preserving the Dual Banking System*, 58 FORDHAM L. REV. 1133, 1156 (1990).

22. *Id.* at 1157–58.

well as the ability to increasingly engage in additional financial services such as insurance and securities brokerage.[23]

The dual banking system also provides protection against oppressive regulation. Bankers may feel that their regulator is overbearing and that regulatory mandates are negatively affecting their ability to manage their bank. Bank executives sometimes argue that regulators cross the line between regulation of the institution and management of the institution. While these concerns may at times provide a scenario for a charter conversion, in most cases it is likely that the regulator, be it state or national, was addressing legitimate "safety and soundness" concerns, and the banker will not find a safe haven with an alternative regulator. Further, though many state banking departments regulate multi-billion dollar banking companies, not all state agencies have experience with such large and complex institutions. For this reason, some large interstate or international companies may opt for OCC regulation, believing the national regulator will be more understanding of their operational issues and challenges. Alternatively, some bankers prefer a more provincial regulatory approach, expecting local regulators to be more sympathetic to, and familiar with, local economic issues and idiosyncrasies.[24]

II. CURRENT STATE PARITY LAWS

A. Near-Unanimous Adoption of Some Parity Provision, Commonly in Conjunction with "Incidental and Proper" Clauses

Nearly every state has enacted some form of parity provision. In fact, only two states, Iowa and North Carolina, have not.[25] Of the forty-eight states that do have state bank parity statutes, the vast majority of their banking codes also include some type of "incidental and proper" provisions that can also serve to expand upon the powers that are specifically enumerated by the legislatures.[26] These clauses have been subject to both narrow and broad interpretations, not unlike the application of the "incidental powers" clause contained in the National Bank Act.[27] The Chief Counsel of the OCC, Julie L. Williams, has proclaimed a broad interpretation of this clause, describing the "business of banking" authorized for national bank charters as "an evolving activity that could be

---

23. *Id.* at 1158.

24. *See* Heidi Mandanis Schooner, *Recent Challenges to the Persistent Dual Banking System*, 41 ST. LOUIS U. L.J. 263, 273 (1996); *see also* Michael L. Stevens, Vice President of Education for the Conference of State Bank Supervisors, Editorial: *Examiners Get Thorough Training for a Bank Career*, AM. BANKER, Jan, 25, 2002, at 16.

25. *See* Survey Results, *infra*; *see also* Conference of State Bank Supervisors—2000 Profile of State-Chartered Banking, Table—Wildcard Authority & Parity Statutes—Part I (on file with author).

26. Survey Results, *infra*.

27. *See* Julie L. Williams & Mark P. Jacobsen, *The Business of Banking: Looking to the Future*, 50 BUS. LAW. 783, 786 (1995); *see also* the National Banking Act, 12 U.S.C. § 24 (2000).

responsive to developments in the financial marketplace and the needs of banks' customers."[28] The breadth of the OCC's interpretation of national bank powers is further expressed in her statement that "[t]he incidental powers granted national banks to conduct activities that are 'incidental' to banking are a separate source of authority to undertake activities that are inherently not part of the business of banking."[29] The OCC periodically updates its list of permissible activities. The most recent issuance is dated February 2001.[30]

Williams' confidence in making such statements stems from the United States Supreme Court decision in *NationsBank of North Carolina v. Variable Annuity Life Insurance*.[31] The Court, in considering "whether national banks may serve as agents in the sale of annuities," supported the OCC's determination that this activity was "incidental to the 'business of banking.'"[32] The opinion reiterated the Court's prior holding that "[i]t is settled that courts should give great weight to any reasonable construction of a regulatory statute adopted by the agency charged with the enforcement of that statute" and stated that the OCC "warrants the invocation of this principle."[33] The Court further held that "the 'business of banking' is not limited to the enumerated powers" and that the OCC "has discretion to authorize activities beyond those specifically enumerated."[34]

While state bank parity laws are generally thought to provide a level playing field between state and national banks, some of the states have expanded the idea of parity beyond federal institutions. For example, the Michigan parity provision provides its state-chartered banks with powers granted to all financial service providers chartered not just by the federal government, but also by any other state or political subdivision.[35] Further broadening this parity provision is the fact that the term "financial service providers" is not defined. Georgia's parity statute also goes beyond federal financial institutions, and includes "others providing financial services in this state existing under the laws of the United States, other states, or foreign governments."[36]

It was noted earlier that neither Iowa nor North Carolina contain parity provisions in their banking codes. While neither of these states have statutes that specifically provide for parity with respect to national banks, representatives

---

28. Julie L. Williams & James F.E. Gillespie, Jr., *The Business of Banking: Looking to the Future—Part II*, 52 BUS. LAW. 1279, 1281-82 (1997).

29. *Id.* at 1282.

30. Office of the Comptroller of the Currency, Activities Permissible for a National Bank (Feb. 2001), *available at* http://www.occ.treas.gov. The preamble to this most recent issuance states: "The business of banking is an evolving concept and the permissible activities of [national banks] similarly evolve over time. Accordingly, this list is not exclusive." *Id.*

31. 513 U.S. 251 (1995).

32. *Id.* at 254.

33. *Id.* at 256 (quoting Clarke v. Sec. Indus. Ass'n, 479 U.S. 388, 403-04 (1987) and Inv. Co. Inst. v. Camp, 401 U.S. 617, 626-27 (1971)).

34. *Id.* at 258.

35. MICH. COMP. LAWS ANN. § 487.14101(2)(b) (1998 & Supp. 2001).

36. GA. CODE ANN. § 7-1-61(a)(1) (1997 & Supp. 2001).

from both states expressed strong opinions that other available legislative provisions serve to provide their respective state banks with all necessary and desired powers.[37] Specifically, Iowa has two provisions that are used in lieu of a parity provision. The first provides that state banks "have and exercise all powers necessary and proper to effect any or all of the purposes for which the state bank is organized."[38] The second and more readily invoked provision provides that Iowa banks may exercise "[a]ll other powers determined by the superintendent to be appropriate for a state bank."[39] According to Donald G. Senneff, Assistant Attorney General and General Counsel to the Iowa Division of Banking, their agency prefers this approach to ensuring competitive parity for two reasons: 1) the state's enumerated powers already provide the ability to engage in the majority of desired activities, and 2) the regulators wanted to avoid conflicts with state laws. He expressed a concern that, in effect, delegating authority to Congress or the OCC could be viewed as a "slap in the face" to the Iowa legislature.[40]

North Carolina statutes provide an even broader powers provision. The enumerated powers list in the "General Powers" article of the state banking code is prefaced with this introduction: "In addition to the powers conferred by law upon private corporations, banks shall have the power . . . ."[41] L. McNeil Chestnut, Special Deputy Attorney General of the North Carolina Department of Justice, points out that this broad provision precludes the need for a parity clause. He notes that, while banks are not required to seek approval prior to exercising new powers, most do so.[42] The North Carolina Commissioner of Banks, in its regulation of the institutions' "safety and soundness," retains the power to oversee, and if necessary, terminate powers or practices deemed unsafe.[43] On an annual basis, the Commissioner publishes a report detailing the various powers engaged in by state banks.[44]

---

37. *See* Survey Results, *infra*; Telephone Interview with Donald G. Senneff, Assistant Attorney General and General Counsel, Iowa Division of Banking (Feb. 5, 2002) [hereinafter Senneff Telephone Interview]; Telephone Interview with L. McNeil Chestnut, Special Deputy Attorney General, North Carolina Department of Justice (Feb. 5, 2002) [hereinafter Chestnut Telephone Interview].

38. IOWA CODE ANN. § 524.801(10) (2001).

39. *Id.* § 524.801(14).

40. *See* Senneff Telephone Interview, *supra* note 37.

41. N.C. GEN. STAT. § 53-43 (1999 & Supp. 2001).

42. *See* Chestnut Telephone Interview, *supra* note 37.

43. N.C. GEN. STAT. § 53-104 (1999 & Supp. 2001).

44. *See* State of North Carolina—Commissioner of Banks, 2000 Survey of Revenue Producing Services, *available at* http://www.banking.state.nc.us/forms/banks/20revnsv.pdf (last visited Jan. 31, 2003); *see also* State of Illinois—Office of Banks and Real Estate, Comparison of Powers of Illinois State Commercial Banks and Savings Banks with Powers of Federal Savings Associations and National Banks, *available at* http://www.obre.state.il.us/CPT/COMCL/POSB/TBLCOM/HTM (last visited Jan. 31, 2003).

B. Most Parity Laws Require Some Type of Notice or Approval

Of the forty-eight states with parity laws, thirty-two require the state bank regulatory agency to approve the specific powers before the bank may engage in them.[45] This authority is most commonly vested in the agency's chief executive, and less often, in the agency board.[46] Another eight states, while not specifically requiring approval, provide for notification by the bank, and allow the banking agency to disapprove the practice within a short period of time—generally thirty to sixty days.[47] While not technically an approval process, the effective results can be the same. In another seven states, the state banks may automatically exercise the power held by national banks within their states.[48] In the remaining state (Kentucky), sometimes the power is automatically extended, based on the condition of the bank.[49] Specifically Kentucky, banks with "CAMELS"[50] ratings of 1 or 2 may exercise parity rights without seeking approval.[51]

---

45. Survey Results, *infra*.

46. *Id.*

47. *Id.*; *see also* Warsame v. State, 659 A.2d 1271, 1272 (Md. 1995); State v. Union Tank Car Co., 439 So.2d 377 (La. 1983); State v. Thompson, 627 S.W.2d 298 (Mo. 1982). In *Warsame*, a Maryland appellate court examined a state narcotics law, stating that "[a]ny new substance which is designated . . . under federal law shall be similarly controlled . . . unless the Department objects" and further noting that the state agency's ability to "object" to the incorporation of federal provisions into a state statute supported the constitutionality of a Maryland state law. 659 A.2d at 1273. Generally this line of cases points to the need for the legislation to both require state agency approval (or lack of objection) and provide defined criteria for consideration.

48. Survey Results, *infra*. In Nebraska, one of the states that allows for the adoption of national bank powers without the requirement for state agency review, the powers are limited to those available to national banks at the time of the annual update of state law. *See* NEB. REV. STAT. § 8-1,140 (1997 & Supp. 2001). Thus, the Nebraska law is not prospective in nature, and only references existing federal law. In essence then, the Nebraska legislature is arguably only choosing to incorporate federal language by reference, rather than drafting separate state language, for powers that it has deemed appropriate for Nebraska banks.

49. Survey Results, *infra*.

50. *See* Federal Deposit Insurance Corporation, Keeping the Promise: Recommendations for Deposit Insurance Reform, n.4 (Apr. 2001), *at* http://www.fdic.gov. This document describes the components of the CAMELS rating system to include an analysis of capital adequacy, asset quality, management, earnings, liquidity, and sensitivity to market risk. Each of these components is rated from one (best) through five (worst), and a composite score is awarded for each institution. CAMELS ratings are generated from on-site examinations of the institutions, generally on an eighteen-month cycle. State banking agencies and the OCC, as the banks' chartering authorities, utilize a range of corrective actions in attempts to rehabilitate troubled institutions. These actions begin with requiring resolutions of the bank's board to make certain changes and can continue if necessary to the closing of the bank. *See, e.g.*, Office of the Comptroller of the Currency, An Examiner's Guide to Problem Bank Identification, Rehabilitation, and Resolution 27-61 (Jan. 2001), *at* http://www.occ.treas.gov/prbbnkgd.pdf.

51. KY. REV. STAT. ANN. § 287.102(2) (1998 & Supp. 2001).

Thus, in approximately eighty-three percent of instances, the state banking department retains the right to either deny, or disapprove, the desired activity. In essence, these state executive branch agencies have been delegated the power by their state legislatures to determine when federal bank powers should be extended to state banks. In the remaining seventeen of the states, the state legislatures have, under most circumstances, and likely unknowingly or inadvertently, delegated this authority to either Congress, or to the OCC, acting through its interpretation of the National Banking Act.

It is important to note that in twenty of the states that provide for agency approval, the statute calls for extension of the powers by either rulemaking or regulation.[52] The implications of this requirement, and its relevance with respect to the constitutional question of legislative delegation, is further considered in Part III of this Note.

C. Many Provide No Specific Guidance for Approval

Fifteen of the forty states that empower their banking agencies (either always or sometimes) with the authority to deny or disapprove parity requests contain no specific criteria for the decision-making.[53] In essence, the determination is left to the discretion of the state regulator, heightening the constitutional question.[54] Twelve states require a determination that the new power, if granted, will not threaten the "safety and soundness" of the institution.[55] Another eleven of the states, in recognition of the competition between state and national charters, require a consideration of the resulting effect on bank competition and the dual banking system if the power is not extended to state banks.[56] The remaining two states consider the contemplated power's consistency with the state banking code, and the general public interest, in determining whether to allow the practice.[57] The presence or absence of such consideration criteria, and the nature and extent of the criteria, is further discussed in Part III's analysis of the constitutional question.

---

52. Survey Results, *infra*.

53. *Id.*

54. Hans. A. Linde, *Structure and Terms of Consent: Delegation, Discretion, Separation of Powers, Representation, Participation, Accountability?*, 20 CARDOZO L. REV. 823, 850 (1999). Linde states that "[u]nconditional delegation of open-ended lawmaking power to a single executive, elected or not, amounts to legislative abdication. It is the essence of modern dictatorships and incompatible with a republican form of government." *Id.* (footnote omitted).

55. Survey Results, *infra*.

56. *Id.* As discussed in Part IV *infra*, these concepts of competition and parity raise a somewhat circular philosophical issue. Namely, state banks, through parity provisions, seek the powers granted to national banks, yet the more homogenized the charters become, the less significance is attached to the inherent characteristics of the dual banking system.

57. *Id.*

D. Most Parity Provisions Override Even Specific State Law Prohibitions

In thirty-five states, if the parity law provisions are met, the federal law preempts even state laws that specifically prohibit particular powers or products.[58] Eight other state parity laws contain only minor exceptions to this blanket preemption.[59] Thus, in only five states did the legislature limit the parity law provisions to allow only for additional powers that are consistent with, and/or not prohibited by, existing state law.[60] These findings appear to represent a significant departure from a regulatory environment that has historically only allowed financial institutions to exercise powers that were specifically enumerated in state law.[61]

The wording used in the various states' parity provisions, with respect to powers otherwise prohibited by state law, varies significantly. In some states, such as Alabama, the intent of the legislature is very clear.[62] The parity clause and the provision overriding contrary state law are contained within the same paragraph.[63] Further, the wording, "[t]he provisions of this section shall take priority over, and be given effect over, any other general or specific provisions of the Alabama law relating to banking to the contrary" leaves little room for debate.[64] Other state statutes are less specific, but just as clear by introducing the parity provision with a phrase such as "[n]otwithstanding other provisions of state law . . . ."[65]

Other statutes are not constructed with such clarity. Indiana's parity provision states, in part: "A bank that intends to exercise any rights and privileges that are: (1) granted to national banks; but (2) *not authorized* for banks under the Indiana Code (except for this section) . . . shall submit a letter to the department describing . . . the requested rights and privileges . . . that the bank intends to exercise."[66] The question is whether the phrase "not authorized" was intended, and should be read, to include "prohibited." Certainly all prohibited powers are not authorized, but it is less clear that the phrase "not authorized" was intended to include previously prohibited powers. In other words, it is clear the provision is meant to grant a power to state banks that had not previously been addressed in state law. It is less clear that this was intended to reach issues that were previously addressed, and proscribed, by the legislature. As this question arises as a product of the unique interplay of dual regulation/legislation, specifically enumerated powers, parity provisions, and the particular wording of

---

58. *Id.*

59. *Id.*

60. *Id.* These states are Arizona, Colorado, Massachusetts, Minnesota, and Montana.

61. Johnson, *supra* note 2, at 357.

62. ALA. CODE § 5-5A-18.1 (1975 & Supp. 2001).

63. *Id.*

64. *Id.*

65. *See, e.g.*, ALASKA STAT. § 06.01.020 (2001); CAL. FIN. CODE § 753(b)(1) (1999 & Supp. 2002); IDAHO CODE § 26-1101(3) (2000 & Supp. 2001); 205 ILL. COMP. STAT. ANN. 5/5(11).

66. IND. CODE § 28-1-11-3.2(b) (1998 & Supp. 2001) (emphasis added).

the Indiana Code, there is no definitive answer to this question. Subsection (f), following four paragraphs later, provides additional guidance: "The exercise of rights and privileges by a bank in compliance with and in the manner authorized by this section is not a violation of any provision of the Indiana Code."[67] While this wording appears to reach previously prohibited powers, a simple phrase such as "notwithstanding any other state law" would have been clearer.

In any event, the delegation of authority that allows state banks to engage in a power, or offer a product, that was previously specifically prohibited by state law, appears to represent the strongest case for the position that at least some state bank parity laws represent an unconstitutional abdication of lawmaking responsibility by state legislators.

E. Parity Provisions Have Been Used to Extend a Variety of Bank Powers

Through the years, parity provisions have been invoked to provide a wide variety of previously unauthorized powers to state banks.[68] Many have afforded additional options or further definitions for core bank products and services. For example, they have provided for amendments and/or additions to the types of lending activities in which banks may engage, as well as adjustments to the calculation of the banks' legal lending limits.[69] The lists of statutorily acceptable investment securities for bank purchase have also been expanded,[70] and some states have expanded the banks' ability to purchase bank-owned life insurance.[71] The provisions have also been used to expand the banks' ability to invest in subsidiaries, and to expand the powers in which bank subsidiaries may engage.[72] These types of state statute amendments that extend federal powers to state banks would be viewed as falling within a historically narrow interpretation of powers that are incidental to the business of banking.

The parity provisions have also been used to expand powers that were traditionally outside of the scope of the business of banking.[73] Common among these are additional powers relative to insurance sales, securities brokerage, and investment advice. Additional powers not previously enumerated by state legislatures include courier services, travel agency services, real estate holdings and leasing, tax preparation service, title insurance powers, and the ability to purchase Federal Home Loan Bank Stock.[74] While not necessarily incidental to the business of banking, these powers and products have come to be considered complimentary to the basic financial services previously offered by banks, thus satisfying the parity provisions of many states, and allowing for greater equality

---

67. *Id.*
68. *See* Survey Results, *infra*.
69. *Id.*
70. *Id.*
71. *Id.*
72. *Id.*
73. *Id.*
74. *Id.*

with the national charter.

III. CONSTITUTIONALITY OF STATE BANK PARITY LAWS

A. Concerns for Abdicating Lawmaking to the Federal Legislature and/or Regulators

Despite the fact that "all enactments enjoy a strong presumption of constitutionality,"[75] an argument can be made that state bank parity laws, which extend the powers of national banks to state-chartered banks, represent an unconstitutional delegation of lawmaking powers by a state legislature to Congress. The primary argument is that parity laws go well beyond the generally accepted practice of the incorporation of certain federal language, and amount to the delegation of authority that is significant and integral to the states' regulation of their financial institutions. In practice, this question of delegation reaches even further since it is the OCC, rather than Congress, whose interpretations and actions often expand national bank powers. This concern is heightened when the powers are extended automatically, without the satisfaction of specified criteria, and heightened further when they override a theretofore-specific prohibition in state law.

The United States Supreme Court, regarding the constitutionality of congressional delegation of legislative powers, considers whether Congress "has attempted to abdicate, or to transfer to others, the essential legislative functions with which it is vested by the Constitution."[76] While the applicability of the analysis of congressional delegation to the issue of state legislative delegation has been debated, in the absence of a Supreme Court opinion directly on point, such precedent remains "entitled to respectful attention and may be relied upon."[77] The questions remain, "what constitutes 'essential legislative functions?'" and further, "what type of legislative delegation amounts to an abdication of these functions?" When state legislators have traditionally provided specifically enumerated bank powers, it can at least be argued that the maintenance of this list is an "essential legislative function." Further, the delegation of the ability to expand this list of approved powers to include activities that were previously specifically prohibited by prior legislation can be argued to represent an abdication of legislative function and responsibility.

Many of the state supreme court decisions that have considered the

---

75. State v. Gill, 584 N.E.2d 1200, 1201 (Ohio 1992). *See also* Indep. Cmty. Bankers Ass'n of S.D. v. State By & Through Meierhenry, 346 N.W.2d 737, 739 (S.D. 1984) (stating that "[a]ny legislative act is accorded a presumption in favor of constitutionality and that presumption is not overcome until the unconstitutionality of the act is clearly and unmistakenly shown and there is no reasonable doubt that it violates fundamental constitutional principles.").

76. Currin v. Wallace, 306 U.S. 1, 15 (1939).

77. Devlin, *supra* note 7, at 1220 (cautioning against consideration of federal precedent when the state constitutional provisions were uniquely structured, reflecting local history or culture, and not consistent with typical constitutional language).

constitutional question of whether legislative delegation by state legislators to Congress or federal agencies was constitutional have involved the simple incorporation of some federal legislative language in state law.[78] In such cases, the courts have generally upheld the constitutionality of the state law on the grounds that reference to a federal definition (in this case, in the state's revenue code) "does not constitute an unconstitutional delegation of legislative power where the prospective recognition is only incidental to the administration of the statute . . . and not likely to frustrate the purpose of the statute."[79] Similar results have been reached when the deference to federal law was only with respect to the definition of terms, such as "bank holding company."[80] This type of language incorporation exercised by state legislatures can serve to ensure consistency and can reduce misunderstandings, especially in areas of interstate commerce. Adoption of certain standard definitions and terminologies can promote efficiency without sacrificing or abdicating state lawmaking powers or state autonomy.

However, while from a practical standpoint it can be argued that it is both expedient and convenient to tie state bank powers to federal powers, it also introduces a very slippery slope. In rebuking what it found to be an unconstitutional delegation of state lawmaking power to the federal government in a labor contract matter, the court in *DeAgostina v. Parkshire Ridge Amusements* stated that "[a]ssuming . . . the means adopted is more practical and convenient than the establishment by the state of its own code authorities modeled after the federal system, that alone presents no justification for what has been done" and warned "if the state's power to delegate governmental functions to a foreign agency is sanctioned, there can be no legitimate limits to its exercise."[81] This statement recognizes potential pitfalls of the temptation to enact state legislation that, by incorporating substantive federal law, puts the value of convenience and timeliness above the need to address the specific

---

78. First Fed. Sav. & Loan Ass'n of New Haven v. Connelly, 115 A.2d 455, 492 (Conn. 1955); State v. Johnson, 173 N.W.2d 894, 895 (S.D. 1970).

79. Miller v. State Dept. of Treasury, 188 N.W.2d 795, 808 (Mich. 1971) (holding that the simple adoption of the federal calculation of "taxable income" did not amount to an abdication of lawmaking by the state legislators).

80. *Indep. Cmty. Bankers Ass'n*, 346 N.W.2d at 743. In the case, the definition used by the court was two and one-half pages long, arguably reason enough to incorporate by reference rather than spell out similar wording in the state law. *Id.* at 744. Interestingly, the plaintiff in this case also challenged the constitutionality of the delegation because the state law referred to "the Bank Holding Company Act of 1956, *as amended*." *Id.* at 743 (emphasis added). That court, citing *State v. Julson*, 202 N.W.2d 145 (N.D. 1972), distinguished between the phrases "and all amendments" and "as amended," and declared that "as amended" referred to the past tense, meaning that it included amendments to the Bank Holding Company Act enacted prior to the incorporation of this definition into state law. *Indep. Cmty. Bankers Ass'n*, 346 N.W.2d at 743-44. Thus the incorporation of a federal law, *as amended*, into a state statute, is not necessarily a prospective delegation, and may be limited to then-existing federal law.

81. 278 N.Y.S. 622, 629-30 (N.Y. Sup. 1935).

legislative issues of the individual state.

The prospective nature of the delegation afforded by parity provisions provides additional pause as in most cases the state legislature is providing for the adoption of powers not yet enacted by Congress. This was addressed by the court in *Independent Community Bankers Ass'n* when it stated, "[s]tatutes adopting laws or regulations of . . . the federal government . . . effective at the time of adoption are valid, but attempted adoption of future laws, rules or regulations of . . . the federal government . . . generally have been held unconstitutional."[82] The Washington Supreme Court echoed this position in *State v. Dougall* when it declared a state narcotics law unconstitutional as it permitted "future federal designation . . . by means of Board inaction or acquiescence."[83] The vast majority of the states' bank parity laws provide for true prospective delegation, as they do not limit the provisions only to existing federal legislation or regulations, thus heightening the constitutional question.[84]

A further concern in bank regulation exists due to the aggressive and arguably liberal interpretations of national bank powers being extolled by the OCC. In effect, such liberal interpretations extend the abdication issue one step further—from the federal legislature to an agency of the Department of the Treasury. National bank powers can arise from one or more of three means: clear legislative authority, prescribed rulemaking procedures by the OCC,[85] or through administrative fiat exercised by the Comptroller. These OCC interpretations can result, and arguably have resulted, in the automatic extension of a power to a state bank that was contemplated neither by state nor federal lawmakers. Indication that the OCC has interpreted the National Banking Act in a manner not foreseen by Congress was evidenced in a 1994 congressional reprimand of the OCC for "'inappropriately aggressive' preemption."[86]

The Louisiana Supreme Court, in *State v. Rodriguez*, addressed the constitutionality of a state law that effectively delegated authority to an agency of the federal government.[87] The state law stated, "The secretary of the

---

82. 346 N.W.2d at 744 (quoting Schryver v. Schirmer, 171 N.W.2d 634, 636-37 (S.D. 1969)); *see also Miller*, 188 N.W.2d at 801 ("It is well settled that incorporation by reference of an existing Federal law in a state statute does not render that statute constitutionally infirm").

83. 570 P.2d 135, 138 (Wash. 1977).

84. One exception, as noted *supra*, note 48, is the Nebraska parity provision that extends only federal powers existing at the time of the enactment of state law. Another is the current South Dakota parity statute that seeks to extend parity only to federal "powers and authorities conferred as of January 1, 1999." S.D. CODIFIED LAWS § 51A-2-14.1 (1990 & Supp. 2001). While these two states' laws appear to avoid prospective delegation to federal law, they also appear to minimize the utility of the parity provisions as a means of providing for statutory amendments between state legislative sessions.

85. *See* Office of the Comptroller of the Currency, OCC Standards for Developing Regulations (Nov. 20, 2001) *available at* http://www.occ.treas.gov.

86. Stacy Mitchell, *Rogue Agencies Gut State Banking Laws*, THE NEW RULES, Fall 2001, at 4 (quoting Congress's 1994 reprimand at pp. H6625-H6642 of the Congressional Record).

87. 379 So.2d 1084 (La. 1980).

Department of Health and Human Resources shall add a substance as a controlled dangerous substance if it is classified as a controlled dangerous substance by the Drug Enforcement Administration of the United States government."[88] While at first glance this statement appears to be a delegation to the secretary of the state agency, the word "shall" effectively ties that official's hands. The state law provides for the automatic inclusion of narcotics if designated by the DEA. In striking the law down as unconstitutional, the court said that the legislature may confer powers "upon executive agencies if it supplies adequate standards to execute legislative policy; however, it cannot surrender the legislative power itself to determine what the law shall be."[89] In this criminal case, the prospective nature of the delegation was significant. Once the DEA added the controlled substance to its list, Louisiana likewise added it, and introduced state legislation to incorporate the substance into their criminal code. However, the plaintiff was arrested and charged with possession between the time the substance was added to the agency's list and the time it was legislatively incorporated into state law. Upon these facts, the prosecution was dismissed.[90]

In 1994, the Texas Department of Banking ("TDB") found itself in the unusual position of promoting a revised parity statute in an effort to *limit* the ability of state banks to undertake powers authorized for national banks. The TDB's efforts were due to the existence of a state constitutional provision, added in 1984, that stated, "A state bank . . . notwithstanding any other provision of this section, has the same rights and privileges that are or may be granted to national banks of the United State domiciled in this State."[91] In her testimony before the Texas House of Representatives, former TDB Commissioner Catherine A. Ghiglieri stated the agency's position that "state bank regulation would be chaotic and unpredictable if Section 16(c) is fully self-activating, and would damage the dual banking system."[92] Conversely, she noted that if, instead, the constitutional provision was viewed as "fully permissive, the Legislature through laws, . . . or the Banking Commissioner through opinions or policies would have to authorize the activity before a state bank could exercise a national bank right."[93] Ultimately a parity provision was enacted that provided a means for state banks to "have the same rights and privileges as national banks" while establishing "an orderly system of implementation . . . essential to regulatory control."[94]

---

88. *Id.* at 1085.

89. *Id.* at 1087.

90. *Id.*

91. TEX. CONST. art. XVI, § 16(c).

92. *Proposed Legislation to Modernize the Texas Banking Code of 1943, Supplement to Testimony Presented to the Committee on Investments & Banking, Texas House of Representatives* (Sept. 22, 1994) (written testimony of Catherine A. Ghiglieri) (on file with author).

93. *Id.*

94. *Id.*; TEX. REV. CIV. STAT. ANN. art. 342-3.010 (1973 & Supp. 1996).

B. Concerns for Abdicating Lawmaking to Executive Branch Officials

Parity provisions that delegate the ultimate determination of the extension of federal banking powers to officials within the executive branch of state governments, specifically the state banking departments, raise another constitutional issue. These provisions can serve, in varying degrees according to their wording and parameters, to provide state banking agencies with legislative-type authority. This concern is exacerbated by parity laws that provide little or no criteria to be considered by the agency in determining whether or not to allow for the extensions of bank powers. The lack of established consideration criteria raises the question of the line between administrative and legislative powers. In discussing this issue, the Kansas Supreme Court noted that "[a]dministrative power is the power to administer or enforce a law, as opposed to the legislative power to make a law," and the determination between the administration and the making of law "depends upon the amount of specific standards included within the delegation."[95] The importance of legislative standards for the executive agency was echoed by the high court of New York when it found "no constitutional prohibition against the delegation of power, with reasonable safeguards and standards, to an agency . . . to administer the law as enacted by the legislature."[96] It is necessary, of course, that executive agencies retain sufficient latitude to effectively administer statutes, and it is clearly impossible for state legislative bodies to anticipate all potential ramifications of newly enacted legislation. While this inability to predict all potential ramifications of new statutes presents an argument for legislation lacking in specificity, the New Jersey Supreme Court cautioned that while "exigencies of modern government have increasingly dictated the use of general, rather than minutely detailed standards" in legislation, it is necessary that statutes "provide adequate restraints on the discretion" of the agency.[97]

While delegation of certain administrative duties and interpretations to executive branch agencies is commonplace,[98] the delegation of the types of determinations encompassed by parity provisions is arguably different. Delegation within parity provisions does not involve simply allowing the agencies the authority to draft policies and procedures for the implementation of statutes. Instead, the agencies are given the charge of determining whether or not to, in effect, augment and expand powers that are otherwise, and have historically been, specifically enumerated.

---

95. Citizens' Util. Ratepayer Bd. v. State Corp. Comm'n, 956 P.2d 685, 707 (Kan. 1998). The court went on to say that, with respect to criteria provided to administrative agencies, "the standards only have to be sufficiently reasonable and definite." *Id.* at 711.

96. Boreali v. Axelrod, 517 N.E.2d 1350, 1354 (N.Y. 1987).

97. Roe v. Kervick, 199 A.2d 834, 857 (N.J. 1964).

98. *See* Curry v. State, 649 S.W.2d 833, 835 (Ark. 1983) (noting that "the limitation against the delegation of lawmaking power does not prevent the General Assembly from authorizing boards or commissions to determine facts upon which the law would be put into execution").

An argument in favor of this type of delegation is found in *Citizens*, a case unrelated to banking, where the court noted that a "modern trend, which we ascribe to, is to require less detailed standards and guidance to the administrative agencies in . . . areas of complex social and economic problems."[99] Certainly the regulation of banks can be argued to encompass complex social and economic problems. Further, experienced bank regulators are undoubtedly better qualified than are state legislators to determine which bank powers and products are prudent. These regulators are also much more familiar with the federal powers that might become subject to parity provisions. However, this can be argued to be the case in many legislative matters. Certainly state health officials understand medical matters more fully than legislators. It is in recognition of this fact that our legislators do not draft legislation in vacuums, but instead solicit and consider significant input from industry professionals and community groups. State legislatures cannot simply delegate all lawmaking authority in complex matters in the interests of efficiency and convenience. Though not dealing specifically with the issue of delegation, the United States Supreme Court, in *Immigration and Naturalization Service v. Chadha et al.*, cautioned, "the fact that a given law or procedure is efficient, convenient, and useful in facilitating functions of government, standing alone, will not save it if it is contrary to the Constitution. Convenience and efficiency are not the primary objectives . . . of democratic government."[100]

Even in the cases where the state legislatures have included criteria for consideration by the regulatory agency, some of the criteria tend to be somewhat vague and nebulous. Though the phrase "safety and soundness" might be argued to fit this description, those in the bank regulatory profession view "safety and soundness" as a clear, concrete measure of a bank's operational integrity, performance, and condition. Many well-defined criteria are considered in determining whether a bank is operating in a safe and sound manner.[101] Further, all bank regulators utilize a standard component rating system known as "CAMELS" in assessing the current and future risk associated with banks' operations.[102] Thus, if the ability to adopt additional power through parity is based on the determination of the "safety and soundness" of the activity, well defined measures are available to banking agencies. It is important to note that "safety and soundness" analysis is applied to both the overall condition of the

---

99. 956 P.2d at 711.

100. 462 U.S. 919, 944 (1983); *see also* Royce C. Lamberth, *Reflections on Delegation in the Chevron Era*, 56 FOOD & DRUG J. 11, 13 (2001) (stating, "[o]n the one hand, delegation is a practical necessity for our country; on the other hand, it is at odds with our democratic roots").

101. *See* Sharpe, *supra* note 5, at 236-44 (describing the prescribed standards to be considered in evaluating "safety and soundness," as required by the Federal Deposit Insurance Corporation Improvement Act of 1991, Pub. L. No. 102-242 (1991) (codified at 12 U.S.C. § 1831a (2000)). Included among the factors that must be considered are: internal controls; loan documentation, underwriting, and quality; interest rate risk; asset growth; earnings; capital adequacy, and, other measures deemed appropriate. Sharpe, *supra* note 5, at 236-37.

102. *See supra* note 50.

bank, and to particular activities engaged in by the bank.[103] When a bank engages in an "unsafe and unsound" practice, it is subject to an order from its regulator to "cease and desist" from that practice.[104] When the overall condition of the bank deteriorates and is deemed "unsafe and unsound," this will be reflected in the bank's CAMELS rating, and will prompt appropriate regulatory action.[105] By providing for state banking agency consideration of "safety and soundness" issues prior to granting powers through parity, state legislatures have attempted to ensure that powers are only extended when they represent a prudent banking practice, and only to banks that are in a condition that is conducive to undertaking new powers.

On the other hand, some of the criteria included for consideration in state parity laws, such as "public convenience and necessity," "competitive equality," and "public interest," while relevant, do not provide very specific guidance. Thus, even when such criteria are applicable, determinations by the agency remain effectively discretionary. The level of discretion afforded the agency has been a primary determining factor in state court decisions that have considered this constitutional issue. Courts have looked for specific criteria to be evaluated by the executive branch agency. In *Curry v. State*, the Arkansas Supreme Court found delegation to an administrative agency constitutional because the statute enumerated several criteria to be considered in determining whether federal designations for controlled substances should be incorporated into state actions.[106] In contrast, the Ohio Supreme Court in *State v. Lyman* found an unconstitutional delegation stating, "it is quite clear that no standards . . . were incorporated in the statute, or so far as we can discover, in any other law."[107]

Though the United States Supreme Court has not specifically addressed this constitutional issue as it relates to state legislatures, insight can be gleaned from their decisions regarding similar congressional questions. While in *Panama Refining Co. v. Ryan*, the Court found unconstitutional delegation, stating, "Congress . . . declared no policy[,] . . . established no standard[,] . . . laid down no rule[,] . . . no requirement, no definition of circumstances and conditions" for application,[108] this decision is not consistent with the majority of the Court's opinions. The prevailing position both before and after *Panama* provides Congress more latitude, requiring only that Congress "lay down . . . an intelligible principle to which the person or body authorized to [make rule] is directed to conform . . . ."[109] Further, the Court has determined "[i]t is not necessary that Congress supply administrative officials with a specific formula for their guidance in a field where flexibility and the adaptation of the congressional policy to infinitely variable conditions constitute the essence of the

---

103. *See* Sharpe, *supra* note 5, at 222-23.
104. *Id.* at 223-24.
105. *Id.*
106. 649 S.W.2d 833, 836 (Ark. 1983).
107. 1987 WL 19033, 5 (Ohio Ct. App. 1987).
108. 293 U.S. 388, 430 (1935).
109. Hampton & Co. v. United States, 276 U.S. 394, 409 (1928).

program."[110]

This question of delegation to executive agencies can be of particular concern in certain political environments. In most states, the senior officials of the state banking agencies are subject to political appointments, and often change based on gubernatorial elections.[111] This is especially true of the agency's chief executive officer, but can also extend to chief deputy and general counsel positions. When agencies experience significant executive-level turnover, they can lose valuable "institutional knowledge."[112] The loss of this perspective that was previously provided by such experience and expertise would limit the agency's ability to make prudent decisions that are sensitive to longer-term industry and regulatory concerns.[113] This type of agency turnover at the decision-making level can minimize the argument for providing significant latitude when legislation deals with "areas of complex social and economic problems,"[114] since the assumed expertise and experience may be lacking.

As noted earlier, the fact that many of the states' parity laws provide for extension of federal powers only through administrative regulation or rulemaking may, to a certain extent, obviate concerns arising from the delegation to executive agencies. Certainly this action by the respective state legislatures clearly bolsters the argument that sufficient parameters and standards are in place. And while the rulemaking/regulation process is not entirely consistent from state to state, its requirements for publication, public hearings, and other democratic features at least afford a significant safeguard against the exercise of administrative fiat by the executive branch.[115] The existence of such a rulemaking requirement was cited as persuasive by the Missouri Supreme Court in their finding of constitutional delegation to an administrative agency in *State v. Thompson*.[116] This position was echoed by the supreme courts of both

---

110. Lichter v. United States, 334 U.S. 742, 785 (1948).

111. Conference of State Bank Supervisors, 2000 Profile of State-Chartered Banking, Table—State Bank Supervisors—Part I (on file with author).

112. Sen. George V. Voinovich, *Crisis in the Federal Workforce: Challenges, Strategies, and Opportunities*, 48-OCT. FED. LAW. 30, 31 (2001) (discussing the potential loss of "an unquantifiable wealth of experience," or "institutional knowledge," by the federal government in the next few years).

113. *Id.*

114. Citizens' Util. Ratepayer Bd. v. State Corp. Comm'n, 956 P.2d 685, 707 (Kan. 1998).

115. Arthur Earl Bonfield, *The Federal APA and State Administrative Law*, 72 VA. L. REV. 297, 316 (1986) (noting that the model state administrative law followed the general notice and comment principles as the federal law). The article describes the model law as being "modeled on the representative, political process of the legislative branch of government. In theory, agency rulemaking in a representative, popularly responsible government, should produce the same result as if the action in question has occurred through action of the legislature." *Id.* at 319 (footnotes omitted).

116. 627 S.W.2d 298, 301 (Mo. 1982). This court further noted the importance of the statute's language that listed "eight specific factors expressed as mandatory considerations . . . in making a determination" of whether to add certain substances to its controlled substance schedules, and an

Alabama and Minnesota, in *Ex parte McCurley*[117] and *State v. King*,[118] respectively.

An issue involving the interplay among state bank powers, delegation, and constitutionality recently arose in Georgia, stemming from a 1997 approval by the Georgia Department of Banking and Finance ("GDBF"). At that time, the agency's commissioner, relying on the state's "incidental and proper" provision rather than its parity clause, approved the acquisition of a real estate brokerage business by a state bank.[119] The commissioner's decision was based, in part, on the fact that "federal thrifts, federal credit unions . . . and banks in twenty-five other states" were so authorized.[120] In early 2002, the Georgia Association of Realtors argued that Georgia state banks may not lawfully engage in real estate brokerage services, and cited the Georgia Supreme Court case *Independent Insurance Agents v. Department of Banking & Finance*[121] in support of their position.[122] *Independent Insurance Agents* involved a GDBF decision to authorize a state bank to operate an insurance agency on the basis of Georgia's then-existing "incidental and proper" provision.[123] The provision authorized "all incidental powers as shall be *necessary* to carry on the banking or trust business."[124] Invoking the principle of *ejusdem generis*,[125] the court determined that the insurance brokerage business, though arguably "convenient or useful" for the bank, was not sufficiently "similar in nature" to an "express power" to meet their interpretation of the "necessary" test.[126] Indicating concerns with constitutional issues relative to separation of powers and delegation authority, the court concluded by saying that if insurance powers for banks are needed, "the proper forum to obtain this power is the legislature."[127]

Leslie A. Bechtel, Deputy Commissioner for Legal and Consumer Affairs

---

additional three findings to add a substance to Schedule IV. *Id.* at 302. The importance of the inclusion of these "statutory standards" was emphasized by the court. *Id.* at 302-03.

117. 390 So.2d 25 (Ala. 1980).

118. 257 N.W.2d 693 (Minn. 1977).

119. *See* Telephone Interview with Leslie A. Bechtel, Deputy Commissioner for Legal & Consumer Affairs, Georgia Department of Banking and Finance (Feb. 19, 2002) [hereinafter Bechtel Telephone Interview].

120. Notice accompanying Petition for Declaratory Ruling to Department of Banking & Finance, Georgia Association of Realtors, Georgia Bankers Association, & Community Bankers Association of Georgia (Feb. 8, 2002), *available at* http://www.ganet.org/dbf/dbf.html (on file with author).

121. 285 S.E.2d 535 (Ga. 1982).

122. *See* Bechtel Telephone Interview, *supra* note 119.

123. *Id.* at 536.

124. *Id.* (emphasis added).

125. This term means "[a] canon of construction that when a general word or phrase follows a list of specific persons or things, the general word or phrase will be interpreted to include only persons or things of the same type as those listed." BLACK'S LAW DICTIONARY 535 (7th ed. 1999).

126. *Indep. Ins. Agents*, 285 S.E.2d at 537.

127. *Id.*

with the GDBF, notes that the holding in *Independent Insurance Agents* was argued despite the fact that Georgia's "incidental and proper" provision has changed significantly since 1982.[128] The current law grants "all powers necessary, *convenient*, or *incidental* to effect any and all purposes for which the bank or trust company . . . is organized."[129] The statute further includes such powers needed to "carry on banking, trust, or other activities determined by the commissioner to be *financial in nature or incident or complementary to such financial activities*," [130] clearly broadening the legislative grant of power. Despite the questionable precedent of *Independent Insurance Agents* due to the change in statutory language, and in part to avoid jeopardizing the GDBF's future ability to expand state bank powers, the Department issued a declaratory ruling stating that it would not approve additional real estate brokerage activities until such powers are granted to national banks.[131] In the event national banks are granted this power, the GDBF can approve the extension of real estate brokerage powers through its parity provision as a means of remaining competitive with the national charter.

IV. PHILOSOPHICAL AND PUBLIC POLICY ISSUES

A. States' Interest in Limiting Bank Powers

Banks gather money in the form of customer deposits and invest that money in loans, securities, real estate, etc. These institutions also benefit from the availability of federal deposit insurance, a safety net that assists them in attracting deposits. Because the citizens of each state are placing their trust and deposits with banks, a compelling state concern arises, and every state maintains banking laws and banking agencies for the purpose of promoting prudent banking practices. A primary means of controlling the level of risk inherent in banking has been limitations of permissible activities, products, and services. Thus, the parity issue has significant ramifications on the safety of the banking system. It is essential that parity provisions do not promote "an unproductive competition in laxity" among regulators.[132] Indeed, in recognition of the danger of burgeoning bank powers, and in the wake of the numerous thrift and bank failures of the 1980s and early 1990s, Congress enacted the Federal Deposit Insurance Corporation Improvement Act of 1991 ("FDICIA"), which curtailed the ability of state legislatures and banking departments to expand state bank

---

128. *See* Bechtel Telephone Interview, *supra* note 119.

129. GA. CODE ANN. § 7-1-261(11) (1997 & Supp. 2001) (emphasis added).

130. *Id.* (emphasis added).

131. *See* Bechtel Telephone Interview, *supra* note 119; Declaratory Ruling of the Georgia Department of Banking & Finance (Feb. 13, 2002) *available at* http://www.ganet.org/dbf/dbf.html.

132. Henry N. Butler & Jonathan R. Macey, *The Myth of Competition in the Dual Banking System*, 73 CORNELL L. REV. 677, 680 (1988) (referring to the competition for charters between state and national regulators).

powers.[133] These concerns remain today, and are arguably heightened by subsequent federal legislation that has brought additional competition for banks and resultant increased pressure on bank earnings performance.

Financial institution failures in the early 1980s were in large part due to institutions actively pursuing newly granted powers without the necessary expertise to adequately assess and control risk.[134] Many thrifts entered into the commercial real estate development business, departing from the safer but generally less profitable home mortgage industry.[135] With the advent of new available powers for banks, particularly insurance and securities underwriting, it will be necessary for regulators to ensure that institutions do not exercise these powers without a clear understanding of the risk.[136] Expansion of powers by means of parity provisions in such an environment raises concerns that did not exist prior to this period of unprecedented expansion of bank powers that has come to be known as "financial modernization."[137]

While parity provisions can be viewed as something akin to "emergency legislation," with the potential for legislative review and possible revision during the succeeding legislative session, practice has not borne this out. Most state legislatures do not meet throughout the entire year. Thus while these parity provisions provide for immediate response to federal legislative initiatives, none of the parity provisions require the legislatures to ratify the adopted powers, thereby statutorily adding them to the previously enumerated list of permissible activities.[138] California's law comes the closest, promoting legislative review/action by including a "sunset" provision. The statute provides that "any regulation . . . shall expire . . . on December 31 of the year following the calendar year in which it became effective."[139] Thus, legislative action is necessary if the grant of power is to be permanent. In addition, some of the states' parity laws include procedures to promote legislative review of powers granted through

---

133. Pub. L. No. 102-242 (1991) (codified at 12 U.S.C. § 1831a (2000)).

134. Federal Deposit Insurance Corporation, An Examination of the Banking Crises of the 1980s and Early 1990s 9-10 (1997), *at* http://www.fdic.gov/bank/historical/history/vol1.html (describing legislative initiatives that, in hindsight, were poor public policy, as they focused on deregulating "the product and service powers of thrifts and to a lesser extent of banks . . . generally unaccompanied by actions to restrict the increased risk taking they made possible").

135. *See* Markham, *supra* note 3, at 245.

136. *See generally* Federal Reserve Bank, Chicago Supervision & Regulation Department, FINANCIAL MODERNIZATION—A GUIDE TO THE GRAMM-LEACH-BLILEY ACT (2000) (on file with author). In the report's cover letter dated April 6, 2000, addressed to all Seventh District state member banks and bank holding companies, John J. Wixted, Jr., Senior Vice President of the Federal Reserve Bank of Chicago, summarized the act as permitting "banks, insurance companies, securities firms, and other financial institutions to affiliate under common ownership and offer their customers a complete range of financial services which were previously prohibited." *Id.*

137. A search of Westlaw's Text & Periodicals Combined (TP-ALL) database yielded 389 hits for the term since January 1, 1996. (last viewed Feb. 23, 2002).

138. *See* Butler & Macey, *supra* note 132, at 705; Survey Results, *infra*.

139. CAL. FIN. CODE § 753(c)(4) (1999 & Supp. 2002).

parity. This legislative review is generally accomplished by the submission to the legislature of a summary of parity actions taken by the state banking department during the prior year.[140] In practice, once a bank is granted powers, the subsequent rescission of those powers could result in significant financial hardship on the bank. This is especially true since virtually all of these adopted powers would involve contractual relationships with customers of the bank, and in many cases third-party providers and servicers. For instance, if a bank were granted the right to conduct some type of real estate development, contractual relationships, both long and short term, would arise among potential tenants, architecture firms, construction management firms, real estate brokerage firms, telecommunications firms, utilities, and a myriad of others. In addition, prior to entering a new venture, the bank would have to expand its staff to include employees with particular expertise in the business. These personnel expenses would be just one of many "sunk costs" incurred by the bank in undertaking a new operation. For these reasons, it is simply not reasonable to contemplate that power and authority, once granted to a bank, can readily be rescinded.

B. Necessity and/or Desirability of Consistency in State Laws Nationwide

As noted, in many respects, the parity laws "are neither similar nor uniform."[141] Since the primary responsibility of each of the regulators is to ensure safe and sound bank operations, is this inconsistency illogical? Is it necessary or even desirable for the states to consider adopting model legislation as suggested by Johnson?[142]

Organizational and corporate structures differ from one interstate banking organization to another. From the company's perspective, strategic and operational planning is much more efficient when the company does not have to consider separate and different legislative and regulatory constraints in each state.[143] In an interstate environment, the state charter can only remain competitive with a national charter by providing a "seamless regulatory" environment.[144] Anything less would preclude the necessary "level playing field." Recognition of this principle is evidenced in agreements among state banking agencies made in efforts to streamline interstate regulation:

---

140. *See, e.g.*, KAN. STAT. ANN. § 9-1715 (1991 & Supp. 2000); MASS. GEN. LAWS ANN. ch. 167F, § 2(31) (1997 & Supp. 2001); N.H. REV. STAT. ANN. § 394-A:7(IX) (1998 & Supp. 2001).

141. Johnson, *supra* note 2, at 402.

142. *Id.*

143. *See* Press Release, Conference of State Bank Supervisors, Announcement of the Adoption of the Nationwide Cooperative Agreement (July 25, 1997) (on file with author). The press release hailed the adoption of a "single regulatory point of contact at both state and federal levels," and the provision by the agreement of "increased regulatory certainty and more uniformity."

144. *See* Frequently Asked Questions, Legislative Affairs, Conference of State Bank Supervisors website, *at* http://www.csbs.org/government/legislative (last visited Feb. 17, 2002) (describing the "single point of contact" concept as setting up a "seamless system of supervision for a state chartered bank that wishes to operate interstate.").

> The goals of the parties of this Agreement are to promote a comprehensive nationwide system for safety and soundness of financial institutions, to supervise and examine multi-state banks in cooperation with other states, to foster effective coordination and communication among the parties to facilitate the process of supervision and examination with the least burden to multi-state banks, and to enhance responsiveness to local needs and interests in an interstate banking and branching environment.[145]

The question of "seamless regulation" is of concern for all interstate banking companies, but it is of particular concern for the companies that chose to operate individual state-chartered banks in more than one state. Each bank subsidiary in a multi-bank holding company maintains its own bank charter, and is thus subject to regulation and examination by its individual chartering authority (state agency or the OCC). Consistency of laws, regulations, and regulatory practices is an integral consideration. Absent the type of agreement discussed above, the national bank charter would hold a significant competitive advantage over the state charter.

While states must be cognizant of the operational requirements of banks, and while the survival of the dual banking system requires that national and state charters be competitive, the decision to provide seamless regulation could arguably also lead to the demise of the dual banking system. As noted in Part I, one of the primary advantages of the dual banking system throughout banking history has been the innovation fostered by its competitive nature. This innovation has been a product of the existence of fifty individual state banking codes and state regulators, in addition to the federal banking laws and regulator. The more the state regulators and state banking codes become homogenized, the less justification there is for the continuation of the dual banking system. In a 1994 law review article, Professor Norman Silber argued that too much deference to federal law by state legislatures amounts to "one small step backward for federalism, and a move forward for federalization."[146]

CONCLUSION

Nearly all states have recognized the need for the adoption of state bank parity laws. Failure to do so, or to apply "incidental and proper" provisions in the absence of parity laws, can result in an unlevel playing field with respect to national banks.

Though most states have parity provisions, the various laws are quite

---

145. Article I, Section 2.2 of the Nationwide Cooperative Agreement coordinated by the Conference of State Bank Supervisors to foster regulatory cooperation among state banking agencies. (on file with author).

146. Norman Silber, *Why the U.C.C. Should Not Subordinate Itself to Federal Authority: Imperfect Uniformity, Improper Delegation, and Revised Section 3-102(C)*, 55 U. PITT. L. REV. 441, 444 (1994).

different in their drafting and application. While the need for a model parity law can be debated, it is logical that certain parity law provisions should be consistent given the collective states' interest in promoting safe and sound banking practices. Among the issues that would benefit the system through greater consistency are the extent and applicability of state banking agency review, criteria and procedure for the review, and treatment of theretofore prohibited powers.

While arguments can be made that parity laws result in unconstitutional delegation of lawmaking by state legislatures, contrary positions are also compelling. An argument for the parity provisions is the need for banking organizations to be afforded new powers in a timely and efficient fashion. It is necessary that state-chartered financial institutions remain functionally competitive with their federal counterparts. The opposite position is that, particularly in an era of rapid "financial modernization," a more thoughtful, legislative consideration of the associated risks might be more appropriate. Further, state supreme court cases, while endorsing the constitutionality of simple definitional references to federal law, have not endorsed deference to federal law in substantive matters. And while state high courts have held constitutional legislation that delegates certain authority to executive agencies, this endorsement has generally been based, in part, upon the existence of sufficient criteria for consideration by the state agency. In any case, state bank parity provisions have been in existence for many years and they have yet to be challenged.

State legislatures, and state bank regulators will need to continue to monitor the fine line between providing seamless regulation for interstate banking companies, and the autonomy and independence that have been the hallmark of the dual banking system. Despite the opinions of some commentators that the dual banking system is without merit and results in regulatory duplication,[147] the survival of the system is not presently in doubt. Many bankers will continue to desire to work with a local regulatory presence, irrespective of the debate on the merits of other dual banking system attributes.

---

147. *See generally* Butler & Macey, *supra* note 132.

Table 1 - Summary of State Bank Parity Laws

	Cite	Approval/Review Required?	Approval Guidelines	Approval Procedure	Overrides Specific Prohibitions?	Incidental & Proper Clause?
Alabama	CA 5-5A-18.1	Yes Superintendent	Maintain dual banking and public interest	Silent	Yes	Yes
Alaska	AS 06.01.020	Yes	Public convenience, dual banking	By regulation	Unless protection of law explicit	No
Arizona	ARS 6-184(A)(2)	No	N/A	N/A	No— "Except as prohibited by law. . ."	Yes
Arkansas	ACA 23-47-101(c)	Yes Commissioner	Silent	Silent	Yes	Yes
California	CA FC 753	Yes Commissioner	Silent	By regulation	Yes	No
Colorado	11-2-103(5)	Yes Banking Board	Silent	By rule or regulation	No—"so long as . . . not prohibited"	No
Connecticut	CGS § 36a-250(a)(41)	Disapproval by Commissioner possible	Safety & Soundness	Staff review of notice	Except sale of title insurance	Yes
Delaware	Reg. 5.761.0017	Commissioner notification, may disapprove	Safety & soundness	Review of notification	Yes	Yes
Florida	FS § 655.061	Yes	Dual banking, public interest	By rule or order	Yes	Yes

	Cite	Approval/Review Required?	Approval Guidelines	Approval Procedure	Overrides Specific Prohibitions?	Incidental & Proper Clause?
Georgia	§ 7-1-61	Yes	Safety & soundness	Generally by regulation	Specific prohibitions considered by department	Yes
Hawaii	HRS § 412:5-201	Yes application and approval	Silent	Written approval for first bank	Yes	Yes
Idaho	IS 26-1101(3)	Director notification, may disapprove	Consistency w/ Idaho Bank Act	Review of notification	Yes, unless found inconsistent w/ Bank Act	No
Illinois	205 ILCS 5/5(11)	No	N/A	N/A	Yes	No
Indiana	IC 28-1-11-3.2	Department notified, may disapprove	Safety and soundness	Review of notification	Yes	Yes
Iowa	None	N/A	N/A	N/A	N/A	Yes
Kansas	KSA 9-1715	Yes Commissioner	Institution preservation, dual banking	Review of request	Yes	Yes
Kentucky	287.020 & 287.102	Commissioner, automatic for CAMELS 1 or 2	Silent	Industry-wide finding by Commissio ner	Yes, except for specifically stated exceptions	No
Louisiana	LSA-RS 6:242(C)	Notification to Commissioner, may object	Silent	Review of notification	Yes	Yes
Maine	Title 9-B MRSA§416	No	N/A	N/A	Yes	Yes

	Cite	Approval/Review Required?	Approval Guidelines	Approval Procedure	Overrides Specific Prohibitions?	Incidental & Proper Clause?
Maryland	FI § 5-504	Yes Commissioner	Protect economy, & public interest	Silent	Yes	Yes
Massachusetts	MGL 167F: § 2(31)	Yes Commissioner	Competition, public convenience	By regulation	No—"provided, however, not otherwise prohibited	Yes
Michigan	MCL 487.14101(2)(b)	Yes Commissioner	Safety and soundness	By order or declaratory ruling	Yes	Yes
Minnesota	MS 48.15	Yes Commissioner	Silent	Silent	No—"may not authorize any activity prohibited"	Yes
Mississippi	§ 81-5-1	Yes Banking Board	Silent	By regulation	Yes	Yes
Missouri	362.106.4	Yes Director	Safety and soundness	Review of notification	Yes	Yes
Montana	32-1-362	Yes	Silent	Silent	No—"if not expressly prohibited"	No
Nebraska	§ 8-1,140	No	N/A	N/A	Yes	Yes
Nevada	NRS 662-015	Yes Commissioner	Silent	Silent	Yes	Yes
New Hampshire	NH RSA 394-A	Yes Commissioner	Safety and soundness	3 petitioners, by rulemaking	Yes	No

	Cite	Approval/Review Required?	Approval Guidelines	Approval Procedure	Overrides Specific Prohibitions?	Incidental & Proper Clause?
New Jersey	NJSA 17:9A-24b.1	Yes Commissioner	Safety and soundness	By regulation	Yes	Yes
New Mexico	NMSA 58-1-54 & 58-1-34(A)(2)	Yes Director	Competitive necessity	Silent	Yes	No
New York	BL 14-g & 14-h	Yes Banking Board	Public interest, parity	By rule or regulation	Yes	Yes
North Carolina	None	N/A	N/A	N/A	N/A	No
North Dakota	NDCC 6-03-38	Yes Banking Board	Silent	Silent	Yes	No
Ohio	§ 1121.05 (A)(1)	Yes Superintendent	Silent	By rule	Yes, except for interest rates	Yes
Oklahoma	IV OS § 402(10)	No	N/A	N/A	Yes, unless specifically prohibited	Yes
Oregon	ORS 706.795 & 708A.010	Yes, Director	Director discretion, convenience, competition	Rulemaking sometimes	Yes	Yes
Pennsylvania	7 PS § 201(c)	Notice to Department, may disapprove	Safety and soundness	Review of notification	Yes	Yes
Rhode Island	RIGL § 19-3-1(7)	No	N/A	N/A	Yes	Yes
South Carolina	§ 34-1-110(A)(1)	Yes State Board of FI	Silent	By regulation or operational instructions	Yes	No

	Cite	Approval/Review Required?	Approval Guidelines	Approval Procedure	Overrides Specific Prohibitions?	Incidental & Proper Clause?
South Dakota	51A-2-14.1	Yes Director	Convenience & advantage, competition	By declaratory ruling	Yes	Yes
Tennessee	TN 45-2-601	No, though may subsequently disapprove	Safety & soundness	N/A	Yes	Yes
Texas	TFC §32.009-010	Notification, commissioner may prohibit	Safety & soundness	May adopt rules	Yes, except for enumerated exceptions	Yes
Utah	UCA 7-1-301(3)(a)	Yes Commissioner	Competition, public interest	Silent	Yes	Yes
Vermont	8 VSA § 14106	No	N/A	N/A	Yes	Yes
Virginia	§ 6.1-5.1	Yes Commission	Silent	By regulation	Yes	Yes
Washington	RCW 30.04.215(3)	Yes regarding powers auth for national banks after 8-31-94	Convenience and advantage, competition	Director finding	Yes	Yes
West Virginia	§ 31A-3-2(a)(5)(B) & § 31A-8C-2(a)	Yes Board of Banking and FI	Silent	Generally by rule	Yes	Yes
Wisconsin	221.0322	Yes Division approval	Safety and soundness	By rule	Yes	Yes
Wyoming	W.S. § 13-2-101(a)(xiii)	Yes Commissioner	Silent	By rule	Yes	No

United States Postal Service

Statement of Ownership, Management, and Circulation

1. Publication Title	2. Publication Number	3. Filing Date
Indiana Law Review	0 0 9 0 - 4 1 9 8	9-16-02
4. Issue Frequency	**5. Number of Issues Published Annually**	**6. Annual Subscription Price**
Quarterly	Four	$30.00

7. Complete Mailing Address of Known Office of Publication *(Not printer) (Street, city, county, state, and ZIP+4)*	Contact Person / Telephone
530 W. New York St., Indianapolis, Marion, Indiana 46202-3225	Chris Paynter 317-274-4440

8. Complete Mailing Address of Headquarters or General Business Office of Publisher *(Not printer)*

530 W. New York St., Indianapolis, Marion, Indiana 46202-3225

9. Full Names and Complete Mailing Addresses of Publisher, Editor, and Managing Editor *(Do not leave blank)*

Publisher *(Name and complete mailing address)*
Indiana University School of Law - Indianapolis
530 W. New York St.
Indianapolis, Indiana 46202-3225

Editor *(Name and complete mailing address)*
Kirby Lee
530 W. New York St.
Indianapolis, Indiana 46202-3225

Managing Editor *(Name and complete mailing address)*
Brian K. Groemminger
530 W. New York St.
Indianapolis, Indiana 46202-3225

10. Owner *(Do not leave blank. If the publication is owned by a corporation, give the name and address of the corporation immediately followed by the names and addresses of all stockholders owning or holding 1 percent or more of the total amount of stock. If not owned by a corporation, give the names and addresses of the individual owners. If owned by a partnership or other unincorporated firm, give its name and address as well as those of each individual owner. If the publication is published by a nonprofit organization, give its name and address.)*

Full Name	Complete Mailing Address
Indiana University Board of Trustees	Indiana University - Bloomington Indiana Memorial Union Room 005 Bloomington, Indiana 47405

11. Known Bondholders, Mortgagees, and Other Security Holders Owning or Holding 1 Percent or More of Total Amount of Bonds, Mortgages, or Other Securities. If none, check box → ☒ None

Full Name	Complete Mailing Address
NONE	NONE

12. Tax Status *(For completion by nonprofit organizations authorized to mail at nonprofit rates) (Check one)*
The purpose, function, and nonprofit status of this organization and the exempt status for federal income tax purposes:
☒ Has Not Changed During Preceding 12 Months
☐ Has Changed During Preceding 12 Months *(Publisher must submit explanation of change with this statement)*

13. Publication Title			14. Issue Date for Circulation Data Below	
Indiana Law Review			September 2002	
15. Extent and Nature of Circulation			Average No. Copies Each Issue During Preceding 12 Months	No. Copies of Single Issue Published Nearest to Filing Date
a. Total Number of Copies *(Net press run)*			1110	1365 Issue 35:4
b. Paid and/or Requested Circulation	(1)	Paid/Requested Outside-County Mail Subscriptions Stated on Form 3541. *(Include advertiser's proof and exchange copies)*	732	729
	(2)	Paid In-County Subscriptions Stated on Form 3541 *(Include advertiser's proof and exchange copies)*	-0-	-0-
	(3)	Sales Through Dealers and Carriers, Street Vendors, Counter Sales, and Other Non-USPS Paid Distribution	150	500
	(4)	Other Classes Mailed Through the USPS	10	25
c. Total Paid and/or Requested Circulation *[Sum of 15b. (1), (2),(3),and (4)]*			892	1254
d. Free Distribution by Mail *(Samples, complimentary, and other free)*	(1)	Outside-County as Stated on Form 3541	-0-	-0-
	(2)	In-County as Stated on Form 3541	-0-	-0-
	(3)	Other Classes Mailed Through the USPS	15	10
e. Free Distribution Outside the Mail *(Carriers or other means)*			140	80
f. Total Free Distribution *(Sum of 15d. and 15e.)*			155	90
g. Total Distribution *(Sum of 15c. and 15f)*			1047	1344
h. Copies not Distributed			63	21
i. Total *(Sum of 15g. and h.)*			1110	1365
j. Percent Paid and/or Requested Circulation *(15c. divided by 15g. times 100)*			85%	93%

16. Publication of Statement of Ownership
☒ Publication required. Will be printed in the 36:1 issue of this publication. ☐ Publication not required.

17. Signature and Title of Editor, Publisher, Business Manager, or Owner — Date

Editorial Specialist/Business Manager — 9-16-02

I certify that all information furnished on this form is true and complete. I understand that anyone who furnishes false or misleading information on this form or who omits material or information requested on the form may be subject to criminal sanctions (including fines and imprisonment) and/or civil sanctions (including civil penalties).

Instructions to Publishers

1. Complete and file one copy of this form with your postmaster annually on or before October 1. Keep a copy of the completed form for your records.
2. In cases where the stockholder or security holder is a trustee, include in items 10 and 11 the name of the person or corporation for whom the trustee is acting. Also include the names and addresses of individuals who are stockholders who own or hold 1 percent or more of the total amount of bonds, mortgages, or other securities of the publishing corporation. In item 11, if none, check the box. Use blank sheets if more space is required.
3. Be sure to furnish all circulation information called for in item 15. Free circulation must be shown in items 15d, e, and f.
4. Item 15h., Copies not Distributed, must include (1) newsstand copies originally stated on Form 3541, and returned to the publisher, (2) estimated returns from news agents, and (3), copies for office use, leftovers, spoiled, and all other copies not distributed.
5. If the publication had Periodicals authorization as a general or requester publication, this Statement of Ownership, Management, and Circulation must be published; it must be printed in any issue in October or, if the publication is not published during October, the first issue printed after October.
6. In item 16, indicate the date of the issue in which this Statement of Ownership will be published.
7. Item 17 must be signed.

Failure to file or publish a statement of ownership may lead to suspension of Periodicals authorization.

Be a Contender in the Courtroom

BNA's *Criminal Law Reporter*

Whether you're a defense attorney, prosecutor, judge, or researcher, find out what experienced criminal justice professionals already know: BNA's *Criminal Law Reporter* provides objective, reliable information and a national perspective on the latest cases and issues in criminal law.

- **Stay up to date** with coverage of federal and state court decisions, legislative activities, and administrative developments relating to criminal law.
- **Understand the impact** and implications of current rulings that set precedent, examine new statutes, or tackle controversial issues.
- **Follow U.S. Supreme Court criminal cases** from filing through oral argument to final disposition — you're alerted via e-mail as soon as a major decision is issued.
- **Explore new cases and legislation** involving sentencing, searches and seizures, attorneys' fees, habeas corpus, and much more.

Sharpen your professional edge by including *Criminal Law Reporter* in your day-to-day practice.

FREE TRIALS!
In print for 5 weeks FREE! **CALL 800-372-1033**
On the Web for 15 days FREE!
GO TO www.bna.com/trials/crex511.htm

Visit BNA's Professional Information Center
http://litigationcenter.bna.com

BNA®

Essential information. Expert analysis.